Fundamental work cost of quantum processes

Philippe Faist[1, 2, *] and Renato Renner[1]

[1]Institute for Theoretical Physics, ETH Zurich, 8093 Switzerland
[2]Institute for Quantum Information and Matter, Caltech, Pasadena CA, 91125 USA

(Dated: Sept. 1, 2017)

Information-theoretic approaches provide a promising avenue for extending the laws of thermodynamics to the nano scale. Here, we provide a general fundamental lower limit, valid for systems with an arbitrary Hamiltonian and in contact with any thermodynamic bath, on the work cost for the implementation of any logical process. This limit is given by a new information measure—the coherent relative entropy—which measures information relative to the Gibbs weight of each microstate. Our limit is derived using a general thermodynamic framework which ensures that our results hold as well in the context of other frameworks such as thermal operations. The coherent relative entropy enjoys a collection of natural properties justifying its interpretation as a measure of information, and can be understood as a generalization of a quantum relative entropy difference. As an application, we recover the standard first and second laws of thermodynamics for macroscopic systems as emergent from the microscopic dynamics. Finally, our technique has an impact on understanding the role of the observer in thermodynamics: Our approach may be applied at any level of knowledge, for instance at the microscopic, mesoscopic or macroscopic scales, thus providing a formulation of thermodynamics which is inherently relative to the observer, and enabling a systematic treatment of Maxwell demon-like situations.

I. INTRODUCTION

Thermodynamics enjoys an extraordinary universality—applying to heat engines, chemical reactions, electromagnetic radiation, and even to black holes—which has prompted its further application to small-scale quantum systems. In such a context the information content of a system plays a key role: Landauer's principle states that logically irreversible information processing incurs an unavoidable thermodynamic cost [1]. Landauer's principle has generated a new line of research—information thermodynamics—in which information and thermodynamic entropy are treated on an equal footing [2], in turn providing a resolution to the paradox of Maxwell's demon [3]. Further fueled by the study of quantum thermo-devices and nano-engines [4–11] and the role of information in the foundations of statistical mechanics [12, 13], information thermodynamics has seen significant contributions from a statistical mechanical perspective [14–19] as well as experimental demonstrations of information-driven thermodynamic devices [20, 21].

When studying the thermodynamics of small-scale quantum systems, one needs to define the thermodynamic framework precisely. A customary approach, the resource theory approach, is to investigate the transformations which are possible after imposing a condition in which only specific types of operations are allowed. Such techniques have allowed to understand general conditions under which it is possible to transform one state into another in thermodynamic resource theories [22–29], to study erasure and work extraction in the single-shot regime [30–32], as well as to extend such results to the case where quantum side information is available [33, 34], to situations with multiple thermodynamic reservoirs [35–39], and to the case of a finite bath size [40, 41]. The role of coherence and catalysis has been underscored [42–47], the effect of correlations studied [48–51], and the efficiency of nanoengines investigated [52–55]. Fully quantum fluctuation relations [56] and a second law equality [57] have been derived, and further connections to the recoverability of quantum information exhibited [58]. We refer to Ref. [59] for a more comprehensive review on these approaches to quantum information thermodynamics.

Our main result is a fundamental limit to the work cost of any logical process implemented on a system with any Hamiltonian and in contact with any type of thermodynamic reservoir. Our result accounts for both the necessary changes in the energy level populations in the system and for the thermodynamic cost of resetting any information which needs to be discarded by the logical process. It is valid for a single instance of the process and ignores unlikely events, thus capturing statistical fluctuations of the work cost.

Our thermodynamic framework is specified by imposing a restriction on the operations which can be carried out, along with introducing a battery system allowing to invest resources to overcome this restriction. Our framework is a natural generalization of the setup in Ref. [60] and is closely related to resource theory approaches [23, 25, 28]. More specifically, our framework is based on Gibbs-preserving maps, that is, mappings for which the thermal state is a fixed point. Gibbs-preserving maps are the most tolerant model one can assume, in the sense that if any non-Gibbs-preserving map is allowed for free, arbitrary work can be extracted, rendering the framework trivial. Since in essentially any reasonable thermodynamic framework the allowed free operations preserve the thermal state, we are ensured that our bound still holds in other standard settings such as the framework of thermal operations [25, 28]. We also introduce an information battery, which is a memory register of qubits which are all individually either in a pure state or in a maximally mixed state. The battery allows us to invest pure qubits to enable a logical process which is not Gibbs-preserving.

Our main result is expressed in terms of a new purely information-theoretic quantity, the *coherent relative entropy*. The coherent relative entropy observes several natural properties, such as a data processing inequality, invariance under isometries, and a chain rule, justifying its interpretation as an entropy measure. It is a generalization of both the min- and max- relative

* phfaist@caltech.edu

entropy as well as the conditional min- and max-entropy. In the asymptotic limit of many independent repetitions of the process (the i.i.d. limit), the coherent relative entropy converges to the difference of relative entropies of the input and the output state relative to the input and output Γ operators. Our quantity hence enriches the collection of entropy measures forming the smooth entropy framework [61–63].

In fact, our model may be phrased in purely information-theoretic terms, abstracting out physical notions such as energy or temperature in an operator Γ which may be interpreted as assigning abstract "weights" to individual quantum states. (In the case of a system in contact with a single heat bath, these weights are simply the Gibbs weights, where at inverse temperature β the value $e^{-\beta E}$ is assigned to each energy level of energy E.) Our main result then quantifies how many pure qubits need to be invested, or how many pure qubits may be distilled, while carrying out a specific quantum logical process given as a completely positive, trace-preserving map, subject to the restriction that the implementation must globally preserve the joint Γ operator of the system and the battery. In this picture, the coherent relative entropy intuitively measures the amount of information "forgotten" by the logical process, conditioned on the output of the process, and counted relative to the "weights" encoded in the Γ operator.

We then apply our framework to a large system and consider transitions between thermodynamic states. (Consider, for instance, an isolated gas in a box which is in a microcanonical state, and consider a process which brings the gas to another microcanonical state of different energy and volume.) Remarkably, it turns out that the work cost of any mapping relating two thermodynamic states, as given by the coherent relative entropy, is equal to the difference of a potential evaluated on the input and the output state, regardless of the details of the logical process. For an isolated system, we show that this potential is precisely the thermodynamic entropy. By coupling the system to a piston, capable of furnishing work to the system and of dissipating heat, we recover the standard second law of thermodynamics relating the entropy change of the system to the dissipated heat.

Furthermore, we show how our framework is well suited for describing different observers, accounting for varying levels of knowledge about a quantum system. This feature allows to systematically analyze Maxwell demon-like situations.

The results presented in this paper have been to a large extent reported in the recent Ph.D. thesis of one of the authors [64].

The remainder of the paper is structured as follows. In Section II, we present the general setup in which our results are derived. Our main result, the work cost of any process in contact with any type of reservoir, is presented in Section III. The same section provides a collection of properties of our new entropy measure, a study of a special class of states whose properties allow them to be used as "battery states" storing extracted work, a discussion of how the macroscopic laws of thermodynamics emerge from our microscopic considerations, and an analysis of how to relate different observers in our framework. Section IV concludes with a discussion and an outlook.

II. A FRAMEWORK OF RESTRICTED OPERATIONS

Consider a system S described by a Hamiltonian H_S. In the framework of Gibbs-preserving maps, an operation $\Phi(\cdot)$ is forbidden if it doesn't satisfy $\Phi(e^{-\beta H_S}/Z) = e^{-\beta H_S}/Z$, where β is a given fixed inverse temperature and $Z = \mathrm{tr}[e^{-\beta H_S}]$. That is, $\Phi(\cdot)$ is forbidden if it doesn't preserve the thermal state. Now observe that the condition on $\Phi(\cdot)$ depends on β and H_S only via the thermal state, so we can rewrite the condition in a more general, but abstract, way as follows: An operation $\Phi(\cdot)$ is forbidden if it doesn't preserve some given fixed operator Γ, that is, if it doesn't satisfy $\Phi(\Gamma) = \Gamma$. We trivially recover Gibbs-preserving maps by setting $\Gamma = e^{-\beta H_S}$. We choose to loosen this condition to an operator inequality for technical reasons and for convenience, by requiring only that $\Phi(\Gamma) \leqslant \Gamma$. In fact, we show that any process $\Phi(\cdot)$ satisfying $\Phi(\Gamma) \leqslant \Gamma$ can be dilated to a process $\Phi'(\cdot)$ satisfying $\Phi'(\Gamma) = \Gamma$ on a larger system (Proposition 2 in the Appendix). The advantage of this abstract version of the Gibbs-preserving-maps model is that our framework and its corresponding results may be potentially applied to any setting where a restriction of the form $\Phi(\Gamma) \leqslant \Gamma$ applies, for a given Γ which does not necessarily have to be the Gibbs state. For instance, this facilitates the analysis of settings involving a control system. The way Γ should be defined is determined by which restriction of the form $\Phi(\Gamma) \leqslant \Gamma$ makes sense to require in the particular setting considered.

Our framework is defined in its full generality as follows. To each system S corresponds an operator Γ_S, which may be any positive semidefinite operator. We then define as *free operations* those completely positive, trace-nonincreasing maps $\Phi_{A\to B}$, mapping operators on a system A to operators on another system B, which satisfy

$$\Phi_{A\to B}(\Gamma_A) \leqslant \Gamma_B \; . \tag{1}$$

One may think of the Γ operator as assigning to each state in a certain basis a "weight" characterizing how "useless" it is. As a convention, if Γ_S has eigenvalues equal to zero, then the corresponding eigenstates are considered to be impossible to prepare—these states will never be observed. In the following, a map obeying (1) will be referred to as a Γ-*sub-preserving map*.

As mentioned above, in the case of a system S with Hamiltonian H_S in contact with a single heat bath at inverse temperature β, we essentially recover the usual model of Gibbs-preserving maps by setting $\Gamma = e^{-\beta H_S}$. In the case of multiple conserved charges such as a Hamiltonian H_S, number operator N_S, etc., we recover the relevant Gibbs-preserving maps model by setting $\Gamma = e^{-\beta(H_S - \mu N_S + \ldots)}$ with the corresponding chemical potentials, as expected; furthermore the physical charges don't have to commute [38, 39].

In the context of quantum information thermodynamics, the amount of extracted work can be counted using a variety of equivalent explicit work storage models [2, 25, 26, 42, 60, 65]. Among these, the *information battery* is easily generalized to our abstract setting. An information battery is a register A of n qubits whose Γ operator is $\Gamma_A = \mathbb{1}_A$. (If $\Gamma_A = e^{-\beta H_A}$ for an inverse temperature β and a Hamiltonian H_A, the requirement that $\Gamma_A = \mathbb{1}_A$ is fulfilled by choosing the completely degenerate Hamiltonian $H_A = 0$.) The register starts in a state where λ_1

qubits are maximally mixed and $n - \lambda_1$ qubits are in a pure state. In the final state, we require that λ_2 qubits are maximally mixed and $n - \lambda_2$ are in a pure state. The difference $\lambda = \lambda_1 - \lambda_2$ is the number of pure qubits extracted or "distilled." In this way, we may invest a number of pure qubits in order to enable a process which is not a free operation, or we may try to extract pure qubits from a process which is already a free operation.

Depending on the physical setup, the λ pure battery qubits can be themselves converted explicitly to some physical resource, such as mechanical work. In the case where we have access to a single heat bath at temperature T, a pure qubit can be reversibly converted to and from $kT \ln 2$ work using a Szilárd engine, where k is Boltzmann's constant; thus, a process from which we can extract λ pure qubits is a process from which we can extract $\lambda \cdot kT \ln(2)$ work using the heat bath. More generally, we may replace the information battery entirely by other battery models, such as corresponding generalizations to our framework of the "wit" [28], or the "weight system" [26, 42]. It is known those work storage models are equivalent [28]; this equivalence persists in our framework (Proposition 3 in the Appendix), with an appropriate generalization of the "extracted resource" λ. In the context of several physical conserved charges, and corresponding thermodynamic baths, the pure qubits or "extracted resource" λ may also be stored in different forms of physical batteries, corresponding to different forms of work, such as *chemical work* [38, 39]. Hence, the quantity λ should be thought of as a dimensionless value, expressed in number of qubits, characterizing the "extracted resource value" of the logical process independently of which type of battery is actually used in the implementation, in the same spirit as the *free entropy* of Ref. [38], and bearing some similarity to currencies in general resource theories [66, 67]).

The main question we address may thus be reduced to the following form (Figure 1). Given an input state σ_X, a quantum channel $\mathcal{E}_{X\to X'}$, and operators $\Gamma_X, \Gamma_{X'} \geqslant 0$, the task is to find the maximum number of qubits which can be extracted, or the minimum number of qubits which need to be invested, in order to implement the given channel on the given input state. Note that we require the correlations between the input and the output to be as specified by $\mathcal{E}_{X\to X'}$, a condition which is not equivalent to just requiring that the given input state σ_X is transformed into the given output state $\mathcal{E}_{X\to X'}(\sigma_X)$. Equivalently, we require that the implementation acts exactly as the channel $(\mathcal{E}_{X\to X'} \otimes \mathrm{id}_{R_X})$ on a purification $|\sigma\rangle_{XR_X}$ of the input, where id_{R_X} denotes the identity process on R_X.

Finally, we ignore improbable events with total probability ϵ, which is necessary in order to obtain meaningful physical results [68]. Indeed, in textbook thermodynamics when calculating the work cost of compressing an ideal gas for instance, one ignores the exceedingly unlikely event where all gas particles conspire to hit against the piston at much greater force than on average, a situation which would require more work for the compression but which happens with overwhelmingly negligible probability. For our purposes we may optimize the zero-error work cost over states which are ϵ-approximations of the required state [60], which is a standard approach in quantum information and cryptography [61, 69].

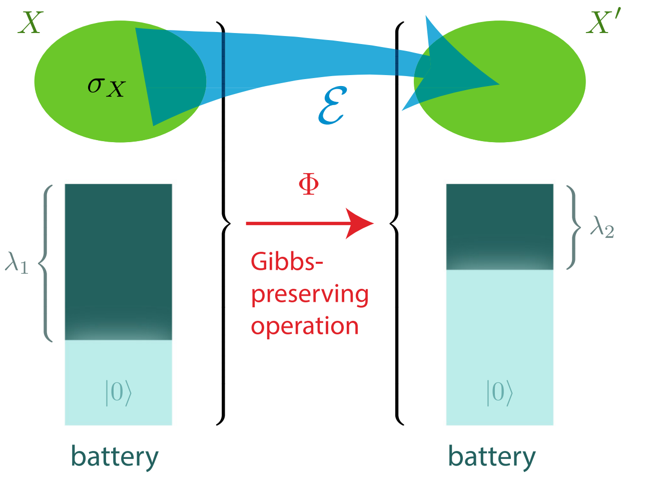

Figure 1. Implementation of a logical process $\mathcal{E}$ (any quantum process) using thermodynamic operations. The process acts on X and has output on X', and is implemented by acting on the system and the battery with a joint Gibbs-preserving operation. The battery starts with depletion state λ_1 and finishes with a depletion state λ_2. The overall extracted work is given by the difference $\lambda_1 - \lambda_2$.

III. RESULTS

A. Fundamental work cost of a quantum process in contact with any bath

Our main result is a fundamental limit on the resource cost required to implement any logical process with thermodynamical operations using any thermodynamic bath. We assume are given a logical process $\mathcal{E}_{X\to X'}$ (defined as any completely positive, trace-preserving map), an input state σ_X, as well as operators $\Gamma_X, \Gamma_{X'}$ corresponding to systems X and X', respectively, as described above and as imposed by the available thermodynamic bath(s) [35, 36, 38, 39]. Then, the optimal implementation of the process with free operations acting on the system and an information battery can extract a number $\lambda_{\mathrm{optimal}}$ of pure qubits, given as

$$\lambda_{\mathrm{optimal}} = \hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \,, \tag{2}$$

where the quantity $\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'})$ is the *coherent relative entropy*. (If $\lambda_{\mathrm{optimal}} < 0$, then the implementation needs to invest at least $-\lambda_{\mathrm{optimal}}$ pure qubits.). Here, R_X is a reference system which purifies the input state as $|\sigma\rangle_{XR_X}$, and the logical process and input state are jointly specified to the coherent relative entropy as the bipartite state $\rho_{X'R_X} = (\mathcal{E}_{X\to X'} \otimes \mathrm{id}_{R_X})(\sigma_{XR_X})$, that is, the output state including correlations with the reference system. The coherent relative entropy measures information relative to the weights represented in Γ_X and $\Gamma_{X'}$, and ignores unlikely events of total probability ϵ (which can be chosen freely).

Recall that the resources required to carry out the process, counted in terms of $\lambda_{\mathrm{optimal}}$ pure qubits, may be converted into physical work. For instance, if we have access to a heat bath at temperature T, we may convert pure qubits into work and vice versa, and thus the work extracted by an optimal implementation of the process is

$$W = kT\ln(2)\cdot \hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})\ . \tag{3}$$

In fact, it is not necessary to implement the process using the information battery at all, and the resources may be directly supplied by a variety of other battery models. The work can even be supplied by a macroscopic piston-like system, as we will see later.

B. The coherent relative entropy and its properties

In the general case, the number of pure qubits which can be extracted from a process, or the negative of the number of pure qubits which need to be invested in order to enable a process, is given by the coherent relative entropy $\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})$. This quantity can be interpreted as a measure of information, as underscored by the collection of properties it satisfies which are natural for a measure of information, and reproduces known results in special cases. We provide an overview here, and refer to Appendix C for the technical details.

a. Definition. The coherent relative entropy is defined as

$$\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) \;=\; \max_{\substack{\mathcal{T}(\Gamma_X)\leqslant 2^{-\lambda}\Gamma_{X'}\\ \mathcal{T}(\sigma_{XR})\approx_\epsilon \rho_{X'R}}} \lambda\ , \tag{4}$$

where the optimization ranges over completely positive, trace-nonincreasing maps $\mathcal{T}_{X\to X'}$, and where the notation $\rho\approx_\epsilon\sigma$ signifies proximity of the quantum states in terms of the *purified distance*, which is related to the ability to distinguish the two states by a measurement [63, 69, 70]. The optimization in (4) is the result of optimizing the battery charge difference in the setup described in Figure 1. In other words, the smooth coherent relative entropy calculates the optimal battery usage in order to implement a given process matrix, up to an accuracy ϵ.

The coherent relative entropy obeys some trivial bounds (Proposition 5 in the Appendix). Specifically,

$$\begin{aligned} &-\log_2\operatorname{tr}\Gamma_X-\log_2\left\|\Gamma_{X'}^{-1}\right\|_\infty+\log_2[1/(1-\epsilon^2)]\\ &\qquad\leqslant\ \hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})\\ &\qquad\leqslant\ \log_2\left\|\Gamma_X^{-1}\right\|_\infty+\log_2\operatorname{tr}\Gamma_{X'}+\log_2[1/(1-\epsilon^2)]\ . \end{aligned} \tag{5}$$

These bounds have a natural interpretation in the context of a single heat bath at inverse temperature $\beta=1/(kT)$. The extracted work may never exceed an amount corresponding to starting in the highest energy level of the system and finishing in the Gibbs state; similarly, it may never be less than the amount corresponding to starting in the Gibbs state and finishing in the highest excited energy level. (A correction is added to account for additional work which can be extracted by exploiting the ϵ accuracy tolerance.)

Under scaling of the Γ operators, the coherent relative entropy simply acquires a constant shift (Proposition 7 in the Appendix): For any $a,b>0$,

$$\begin{aligned} &\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,a\Gamma_X,b\Gamma_{X'})\\ &\qquad=\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})+\log_2\frac{b}{a}\ . \end{aligned} \tag{6}$$

In the case of a single heat bath at inverse temperature $\beta=1/(kT)$, this property simply corresponds to the fact that if the Hamiltonians of the input and output systems are translated by some constant energy shifts, then the difference in the shifts should simply be accounted for in the work cost. Indeed, if $H_X\to H_X+\Delta E_X$ and $H_{X'}\to H_{X'}+\Delta E_{X'}$, then $\Gamma_X\to e^{-\beta\Delta E_X}\Gamma_X$, $\Gamma_{X'}\to e^{-\beta\Delta E_{X'}}\Gamma_{X'}$ and the optimal extracted work of a process, given by $kT\ln(2)$ times the coherent relative entropy, has to be adjusted according to (6) by $kT\ln(2)\log_2(e^{-\beta\Delta E_{X'}}/e^{-\beta\Delta E_X})=\Delta E_X-\Delta E_{X'}$.

b. Recovering known entropy measures. In special cases we recover known results in single-shot quantum thermodynamics, reproducing existing entropy measures from the smooth entropy framework [61, 63]. We refer to Section C 6 in Appendix C for precise definitions of the entropy measures and proofs.

In the case of a system described by a trivial Hamiltonian, it is known that the work cost of erasing a state to a pure state is given by the max-entropy [30], a measure which characterizes data compression or information reconciliation [71]; similarly, preparing a state from a pure state allows to extract an amount of work given by the min-entropy of the state, a measure which characterizes the amount of uniform randomness which can be extracted from the state. These turn out to be special cases of considering the work cost of any arbitrary quantum process for systems with a trivial Hamiltonian [60], which is given by the conditional max-entropy of the discarded information conditioned on the output of the process (Proposition 27):

$$\begin{aligned} &\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\mathbb{1}_X,\mathbb{1}_{X'})\\ &\qquad\approx -H^{\epsilon}_{\max,0}(E\,|\,X')_\rho = H^{\epsilon}_{\min,0}(E\,|\,R_X)_\rho\ , \end{aligned} \tag{7}$$

where $|\rho\rangle_{EX'R_X}$ is a purification of $\rho_{X'R_X}$ and where $H^{\epsilon}_{\max,0}(E\,|\,X')_\rho$ and $H^{\epsilon}_{\min,0}(E\,|\,R_X)_\rho$ are the smooth conditional max-entropy and min-entropy which were introduced in Ref. [61], and are also known as the alternative conditional max-entropy and min-entropy [72]. The approximation in (7) is a technicality related to how to deal with the ϵ parameter (see Appendix); it reduces to an exact equality if $\epsilon=0$.

We recover more known results with an arbitrary Hamiltonian in contact with a heat bath by considering state formation and work extraction of a quantum state [25, 31]. It is known that the work which can be extracted from a quantum state, or which is required to form a quantum state, is given by the min-relative entropy and the max-relative entropy, respectively; these single-shot relative entropies were originally introduced in Ref. [62] and are related to hypothesis testing [73–78]. We show, if the input or output system is trivial, that

$$\hat{D}^{\epsilon}_{X\to\emptyset}(\rho_{R_X}\,\|\,\Gamma_X,1)\approx D_{\min,0}(\rho_X\,\|\,\Gamma_X)\ ; \tag{8a}$$

$$\hat{D}^{\epsilon}_{\emptyset\to X'}(\rho_{X'}\,\|\,1,\Gamma_{X'})\approx -D_{\max}(\rho_{X'}\,\|\,\Gamma_{X'})\ , \tag{8b}$$

matching the previous results. We note that a trivial system as output or input of a process is equivalent to mapping to or from a pure, zero-energy eigenstate; this is because the coherent relative entropy is insensitive to energy eigenstates (or more generally, eigenstates of the Γ operator) which have no overlap with the corresponding input or output state.

c. Data processing inequality and chain rule. The coherent relative entropy satisfies a data processing inequality: If an additional channel is applied on the output, mapping the Gibbs weights to other Gibbs weights, then the coherent relative entropy may only increase. That is, for any channel $\mathcal{F}_{X'\to X''}$,

$$\begin{aligned}&\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})\\ &\qquad\leqslant \hat{D}^{\epsilon}_{X\to X''}(\mathcal{F}_{X'\to X''}(\rho_{X'R_X})\,\|\,\Gamma_X,\mathcal{F}_{X'\to X''}(\Gamma_{X'}))\ . \qquad (9)\end{aligned}$$

Intuitively, this holds because the final state after the application of $\mathcal{F}_{X'\to X''}$ is less valuable as it is closer to the Gibbs state, and hence more work can be extracted by the optimal process realizing the total operation $X\to X''$.

The coherent relative entropy also obeys a natural chain rule: The work extracted during two consecutive processes may only be less than an optimal implementation of the total effective process. We refer to Proposition 19 in the Appendix for a technically precise formulation.

d. Asymptotic equipartition. An important property of the coherent relative entropy is its asymptotic behavior in the limit of many independent copies of the process (known as the *i.i.d. limit*). In this regime, the coherent relative entropy converges to the difference in the quantum relative entropies of the input state to the output state (ϵ is kept small but constant), which is consistent with previous results in quantum thermodynamics [24, 28]:

$$\begin{aligned}&\lim_{n\to\infty}\frac{1}{n}\hat{D}^{\epsilon}_{X^n\to X'^n}(\rho^{\otimes n}_{X'^n R_{X^n}}\,\|\,\Gamma_X^{\otimes n},\Gamma_{X'}^{\otimes n})\\ &\qquad\qquad = D(\sigma_X\,\|\,\Gamma_X)-D(\rho_{X'}\,\|\,\Gamma_{X'})\ , \qquad (10)\end{aligned}$$

recalling that σ_X is the input state of the process and $\rho_{X'}$ the resulting output state. Crucially, the average work cost of performing a process in the i.i.d. regime with Gibbs-preserving operations does not depend on the details of the process, but only on the input and output state, as was already the case for systems described by a trivial Hamiltonian [60].

e. Miscellaneous properties. We show a collection of further properties in Appendix C, including the following: Performing an identity process even on a non-full-rank state cannot be used to extract work with certainty (Proposition 13); the smooth coherent relative entropy can be bounded in both directions as differences of known entropy measures (Section C 7); the coherent relative entropy does not depend on the details of the process if the input state is of the form $\Gamma_X/\mathrm{tr}\,\Gamma_X$ (e.g., a Gibbs state), and reduces in this case to a difference of input and output relative entropies (Proposition 15).

C. Battery states and robustness to smoothing

Previous work has already shown the equivalence of several battery models known in the literature [28], notably the information battery, the work bit (or "wit") [25, 28], and the "weight" system [26, 57]. Our framework allows to make this equivalence manifest, by singling out class of states on any system for which the system can act as a battery. As expected, battery states exhibit the property that they are reversibly interconvertible (as in Ref. [79])—the resources invested in a transition from one battery state to another can be recovered entirely and deterministically by carrying out the reverse transition.

For any system W with a corresponding Γ_W, we consider as battery states those states of the form

$$\tau(P)=\frac{P\Gamma_W P}{\mathrm{tr}(P\Gamma_W)}\ , \qquad (11)$$

where P is a projector such that $[P,\Gamma_W]=0$. In the presence of a single heat bath at inverse temperature β, this class of states includes for instance individual energy eigenstates, or also maximally mixed states on a subspace of an energy eigenspace. We define the value of a particular battery state $\tau(P)$ as

$$\Lambda(\tau(P))=-\log_2\mathrm{tr}(P\Gamma_W)\ . \qquad (12)$$

By requiring the system W to start in such a battery state $\tau(P)$, and requiring the system to end in another such state $\tau(P')$ (corresponding to another projector P' with $[P',\Gamma_W]=0$), then we can show that the system W can act as a battery enabling the exact same state transitions on another system S as an information battery, with $\Lambda(\tau(P'))-\Lambda(\tau(P))=\lambda_1-\lambda_2$ (as long as such states are available in W; cf. Proposition 3 in the Appendix for a precise formulation).

The information battery, the wit as well as the weight system are themselves special cases of this general battery system. Indeed, the states $2^{-\lambda_i}\mathbb{1}_{2^{\lambda_i}}$ of the information battery can be cast in the form (11), with $P=\mathbb{1}_{2^{\lambda_i}}$ since $\Gamma=\mathbb{1}$; the corresponding value of the state is indeed $\Lambda(\tau(P))=-\lambda_i$. Similarly, in the case of the wit and of the weight system, and in the presence of a single heat bath at inverse temperature β such that $\Gamma_W=e^{-\beta H_W}$, the relevant states are energy eigenstates $|E\rangle_W$, whose value is precisely their energy (up to a factor β): $\Lambda(\tau(|E\rangle\langle E|_W))=\beta E$. The equivalence of these models is thereby manifest.

As can be expected, the battery states of the general form $\tau(P)$ are reversibly interconvertible, meaning that for any process which maps $\tau(P)$ to $\tau(P')$ on a system, the coherent relative entropy is equal to the difference $\Lambda(\tau(P))-\Lambda(\tau(P'))$ (Proposition 15).

This general formulation enables us to prove an interesting property of these battery states—they are robust to small imperfections. Indeed, when implementing a process on a system S using a battery W, it makes no difference whether one optimizes over ϵ-approximations of the overall process on the joint system $S\otimes W$, or over ϵ-approximations on S only with no imperfections on the battery state (as the smooth coherent relative entropy is defined above). More precisely, we prove that the smooth coherent relative entropy is exactly the optimal difference in the charge state of the battery, while capturing all implementations which include slight imperfections on the battery for any battery system:

$$\begin{aligned}&\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})\\ &\qquad=\max_{W,W',P_W,P'_{W'},\Phi_{XW\to X'W'}}-\log_2\frac{\mathrm{tr}\,P'_{W'}\Gamma_{W'}}{\mathrm{tr}\,P_W\Gamma_W}\ , \qquad (13)\end{aligned}$$

where the optimization ranges over all battery systems W,W' with corresponding $\Gamma_W,\Gamma_{W'}$, over all battery states corresponding to projectors $P_W,P'_{W'}$ with $[P_W,\Gamma_W]=0$ and $[P'_{W'},\Gamma_{W'}]=$

0, and over all free operations $\Phi_{XW\to X'W'}$ which are an ϵ-approximation of a joint process $XW \to X'W'$ which has a resulting process matrix on the system of interest given by $\rho_{X'R_X}$ and which induces a transition on the battery from $\tau(P_W)$ to $\tau(P'_{W'})$ (see Corollary 40 for technical details).

D. Emergence of macroscopic thermodynamics

We now apply our general framework to the case of macroscopic systems, and recover the standard laws of thermodynamics as emergent.

a. The general mechanism. The macroscopic theory of thermodynamics is recovered when it is possible to single out a class of states which obey a reversible interconversion property. More precisely, suppose there are a class of states $\{\tau^{z_1,z_2,\ldots,z_m}\}$ specified by m parameters $z_1,\ldots,z_m$, and suppose there exists a potential $\Lambda(z_1,\ldots,z_m)$ such that for any pair of states $\tau_X^{z_1,\ldots,z_m}$ and $\tau_{X'}^{z'_1,\ldots,z'_m}$ from this class, we have that for any process matrix $\rho_{X'R_X}$ mapping one state to the other,

$$\ln(2)\cdot \hat{D}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) = \Lambda(z_1,\ldots,z_m) - \Lambda(z'_1,\ldots,z'_m)\,, \quad (14)$$

where the $\ln(2)$ factor serves to change the units of the coherent relative entropy from bits, which is standard in information theory, to nats, which will prove convenient to recover the standard laws of thermodynamics. We call the function $\Lambda(z_1,\ldots,z_m)$ the *natural thermodynamic potential* corresponding to the physics encoded in the Γ operators. In other words, the two states $\tau^{z_1,\ldots,z_m}$ and $\tau^{z'_1,\ldots,z'_m}$ may be reversibly interconverted, as any work invested when going in one direction may be recovered when returning to the initial state, and this irrespective of which precise logical process is effectively carried out during the transition. An obvious choice for these states are states of the same form as the battery states introduced above, which motivates recycling the same symbols τ and Λ. (We have set $\epsilon = 0$ in (14) because smoothing such battery-type states has no significant effect. In more general cases, one might have to consider the condition (14) as valid only in some limit (e.g., a limit of large system sizes), and with some nonzero ϵ taken to vanish in that limit.)

Suppose the parameters are sufficiently well approximated by continuous values. This would typically be the case for a large system such as a macroscopic gas. Consider an infinitesimal change of a state $(z_1,\ldots,z_m) \to (z_1+dz_1,\ldots,z_m+dz_m)$. If there is a free operation which can perform this transition, then necessarily the coherent relative entropy is positive, and hence $\Lambda(z_1+dz_1,\ldots,z_m+dz_m) \leqslant \Lambda(z_1,\ldots,z_m)$. Conversely, if the coherent relative entropy is positive, then there necessarily exists a free operation implementing the said transition. We deduce that the infinitesimal transition $z \to z+dz$ is possible with a free operation if and only if

$$d\Lambda \leqslant 0\,. \quad (15)$$

This condition expresses the macroscopic second law of thermodynamics, as we will see below.

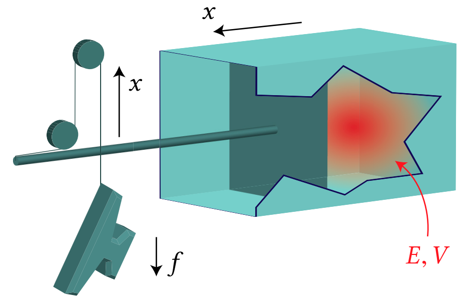

Figure 2. Macroscopic thermodynamics emerges from our framework when singling out a set of states which can be parametrized by continuous parameters to a good approximation, and which can be reversibly interconverted into one another. We consider the case of a textbook thermodynamic gas confined in a box, with a piston capable of furnishing work. In this setting, we recover the usual second law of thermodynamics, $dS \geqslant \delta Q/T$, relating the change in entropy, the dissipated heat and the temperature.

We may define the generalized chemical potentials

$$\mu_i = \left(\frac{\partial\Lambda}{\partial z_i}\right)_{z_1,\ldots,z_{i-1},z_{i+1},\ldots,z_m}\,, \quad (16)$$

where the notation $(\partial f/\partial x)_{y,z}$ denotes the partial derivative with respect to x of a function f, as y and z are kept constant. We may then write the differential of Λ as

$$d\Lambda = \sum \mu_i\,dz_i\,. \quad (17)$$

The generalized potentials μ_i are often directly related to physical properties of the system in question, such as temperature, pressure or chemical potential.

Under external constraints on the variables $z_1,z_2,\ldots,z_m$, we may ask what the "most useless thermodynamic state" compatible with those conditions is. The answer is given by minimizing the potential Λ subject to those constraints—this is a variational principle. For instance, if two systems with natural thermodynamic potentials $\Lambda_1(z_1,\ldots,z_m)$ and $\Lambda_2(z'_1,\ldots,z'_m)$ are put into contact under the constraints that for all i, $z_i + z'_i$ must be kept constant (such as for extensive variables in thermodynamics), then we may write $dz_i = -dz'_i$ and minimize $\Lambda = \Lambda_1 + \Lambda_2$ by requiring that

$$0 = d\Lambda = \sum(\mu_i - \mu'_i)\,dz_i\,, \quad (18)$$

and we see that the minimum is attained when $\mu_i = \mu'_i$. If the system is undergoing suitable thermalizing dynamics, then its evolution will naturally converge towards that point.

b. The textbook thermodynamic gas. We proceed to recover the usual laws of thermodynamics in this fashion for a macroscopic isolated gas S composed of many particles (Figure 2). The Hamiltonian of the gas is denoted by $H^{(V)}$, where the volume V occupied by the gas is a classical parameter of the Hamiltonian which determines for instance the width of a confining potential. We assume for simplicity that the number N of particles constituting the gas is kept at a fixed value throughout, restricting our considerations to the corresponding subspace.

Let us first consider the case of an isolated gas at fixed parameters E, V. In order to apply our framework, we must identify the Γ operator which encodes the relevant restrictions imposed by the physics of our system. Recall that our restriction is meant to explicitly forbid certain types of processes, without worrying whether a non-forbidden operation is achievable. Here, we assume that at fixed E, V the system is isolated and hence evolves unitarily. In particular, at fixed E, V, the operator $P_S^{E,V}$ must be preserved, where $P_S^{E,V}$ is the projector onto the eigenspace of $H^{(V)}$ corresponding to E. Hence, the Γ operator characterizing the gas alone for fixed E, V can be taken as

$$\Gamma_S^{E,V} = P^{E,V} \ . \tag{19}$$

This is compatible with standard considerations in statistical mechanics, which identify the state of the gas in such conditions as the maximally mixed state in the subspace projected onto by $P^{E,V}$ (the Boltzmann microcanonical state), which we denote by $\tau_S^{E,V} = P^{E,V}/\operatorname{tr} P^{E,V}$. Indeed, at fixed E, V on the control system, an allowed transformation may not change this state.

Now we would like to account for changes in E, V. It is convenient to introduce a physical control system C, which plays the following roles: It stores the information about all the controlled external parameters of the state in which the gas was prepared—here, the parameters are E, V; furthermore, it provides the necessary physical constraints on the gas and physical resources necessary for transformations. In our case, the control system includes a piston which confines the gas to a volume V, and is capable of furnishing the energy required to change the state of the gas. For concreteness, we imagine that the piston is balanced by a weight, causing the piston to exert a force f on the gas. The force f may be tuned by varying the weight. The states of the control system are $|e, x\rangle_C$, where e is the energy stored in the control system and x the position of the piston. The energy e is the potential energy of the weight, and must be equal to $e = E_{\text{tot}} - E$ as enforced by total energy conservation, where E_{tot} is the fixed total energy of the joint CS system. Furthermore x determines the volume of the gas as $V = A \cdot x$, where A is the surface of the piston. If the control system were isolated and not coupled to the gas, then the non-forbidden operations on the control system would be those preserving the operator $\Gamma_C^0 = \sum_{e,x} g_{e,x} |e, x\rangle\langle e, x|_C$, where $g_{e,x}$ encodes the relevant physics of the control system: it decreases as either e increases or x increases, meaning that a state $|e, x\rangle_C$ cannot be brought to the state $|e', x\rangle_C$ with $e' > e$ or $|e, x'\rangle_C$ with $x' > x$. In other words, we do not forbid reducing the weight charge or lowering it.

The coupling between the control system and the gas can be enforced with a Γ operator of the form

$$\Gamma_{CS} = \sum_{e,x} g_{e,x} |e, x\rangle\langle e, x|_C \otimes P_S^{E=E_{\text{tot}}-e, V=Ax} \ . \tag{20}$$

If the control system is the state $|e, x\rangle_C$, then any allowed operation must preserve the operator $\Gamma^{E,V}$ for the corresponding $E = E_{\text{tot}} - e$ and $V = Ax$. Furthermore (20) accounts for the physics of the control system itself with the coefficient $g_{e,x}$.

The states $\tau_{CS}^{e,x} = |e, x\rangle\langle e, x|_C \otimes \tau_S^{E=E_{\text{tot}}-e, V=Ax}$ are of the form (11), and hence they are reversibly interconvertable as per (14) and they are a valid class of states for our macroscopic description. The natural thermodynamic potential is

$$\begin{aligned} \Lambda_{CS}(e, x) &= -\ln(g_{e,x} \operatorname{tr} P^{E=E_{\text{tot}}-e, V=Ax}) \\ &= \Lambda_C(e, x) + \Lambda_S(E_{\text{tot}} - e, Ax) \ , \end{aligned} \tag{21}$$

where we have defined $\Lambda_C(e, x) = -\ln g_{e,x}$ and $\Lambda_S(E, V) = -\ln \operatorname{tr} P_S^{E,V}$. Observe that $\operatorname{tr} P_S^{E,V} = \Omega_S(E, V)$ is the microcanonical partition function, and hence $\Lambda_S(E, V)$ is, up to Boltzmann's constant k and a minus sign, the quantity $S(E, V) = k \ln \Omega_S(E, V)$ which is known as the thermodynamic entropy of the gas:

$$\Lambda_S(E, V) = -k^{-1} S(E, V) \ . \tag{22}$$

As the gas is macroscopic, we assume that the parameters E, V are well approximated by continuous variables. It is useful to define the conjugate variables to e, x and E, V via the differentials of Λ_C and Λ_S:

$$d\Lambda_C = \nu_e \, de + \nu_x \, dx \ ; \tag{23a}$$

$$d\Lambda_S = \mu_E \, dE + \mu_V \, dV \ , \tag{23b}$$

with the coupling inducing the relations $dE = -de$ and $dV = A\,dx$. The force f exerted by the piston onto the gas is given by $f = (\partial e/\partial x)_{\Lambda_C}$. Using (23a) we see that $de = \nu_e^{-1}(d\Lambda_C - \nu_x dx)$ and hence $f = -\nu_x/\nu_e$. The thermodynamic work provided by the piston is the mechanical work performed by the weight,

$$\delta W = -f \cdot dx = \frac{\nu_x}{\nu_e} \, dx \ . \tag{24}$$

Since our framework applies to the joint system CS, it follows that any operation mapping two states $\tau_{CS}^{e,x} \to \tau_{CS}^{e+de, x+dx}$ which obeys our global restriction, i.e. which preserves the operator (20), obeys (15) or equivalently $d\Lambda_S \leqslant -d\Lambda_C$, and hence

$$d\Lambda_S \leqslant -\nu_e \, de - \nu_x \, dx = \nu_e \, (dE - \delta W) = \nu_e \, \delta Q \ , \tag{25}$$

where we have defined the change in energy of the gas not due to thermodynamic work as *heat*: $\delta Q = dE - \delta W$.

The temperature of the gas is defined as $T_{\text{gas}} = -(k\mu_E)^{-1} = (\partial S/\partial E)^{-1}$ as in standard textbooks, as the conjugate variable corresponding to entropy. Analogously, we define the temperature T of the piston as the temperature of a gas it would be "in equilibrium" with, in the sense that our variational principle is achieved. The potential Λ_{CS} attains its minimum under the constraints $dE = -de$ and $dV = A \cdot dx$ if $0 = d\Lambda_{CS} = (\mu_E - \nu_e) \, dE + (\mu_V + A^{-1}\nu_x) \, dV$, implying that $\mu_E = \nu_e$ and hence $T = -(k\nu_e)^{-1}$.

We may now write (25) in its more traditional form,

$$dS \geqslant \frac{\delta Q}{T} \ . \tag{26}$$

Our control system is in fact another example of a battery system. Indeed, it can convert another form of a useful resource, mechanical work, into the equivalent of pure qubits for enabling processes on the system, while still working under the relevant global constraints such as conservation of energy.

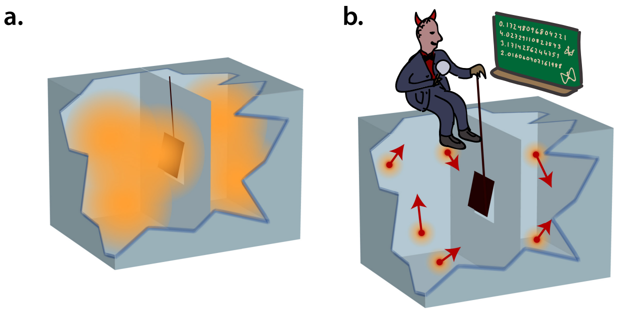

Figure 3. Maxwell's demon concentrates all particles on one side of the box by opening the trap door at appropriate times. **a.** A macroscopic observer describing only the gas sees its entropy decrease, in apparent violation of their idea of the second law of thermodynamics. **b.** The demon observes no entropy change as he correlates his memory (originally in a pure state) with the state of the gas. In doing so he may induce a macroscopic observer into witnessing a violation of a macroscopic second law. If the demon wishes to operate cyclically, he needs to reset his memory register back to a pure state, which costs work according to Landauer's principle [2, 3]; any work he might have extracted using his scheme is paid back at this point.

The thermodynamic gas illustrates a situation in which the first and second laws of thermodynamics are recovered from a conservation law, and as emergent, respectively. Note that the argument can be applied as well to a system with different relevant physical quantities, such as magnetic field and magnetization of a medium.

E. Observers in thermodynamics

In standard thermodynamics, one describes systems from the macroscopic point of view. For instance, thermal equilibrium or the thermodynamic entropy function are usually treated as properties of the system, as defined from the fixed, macroscopic point of view. Yet, a closer look reveals they can be thought of as observer-dependent quantities which can be extended to observers with different amounts of knowledge about the system [33, 34, 80].

This observation is at the core of a modern understanding of most examples surrounding Maxwell's demon. As a concrete example, consider a variant of Maxwell's demon depicted in Figure 3. A gas is enclosed in a box separated into two equal volume compartments, which communicate only through a small trap door controlled by a demon. The demon is able to observe individual particles, and activates the trap door at appropriate times, letting a single particle through each time, in order to concentrate all particles on one side of the box. From a macroscopic perspective, and looking only at the gas, one observes an apparent entropy decrease as the gas now occupies a smaller volume. However, from a microscopic perspective, the demon is essentially transferring entropy from the gas into a memory register, which is initially in a pure state [2, 3]. Consider in more detail the following process: the demon performs a series of c-not gates using the gas degrees of freedom as controls and his memory qubits as targets, which "replicates" the information about the gas particles into his memory. Since this process is unitary, it preserves the joint entropy of the memory and the gas. The result is a classically correlated state between the memory register and the gas. So, *what is the entropy of the gas?* It is now clear that the answer depends on the observer. The macroscopic observer sees the gas with its usual macroscopic thermodynamic entropy, while the demon has engineered a state where the gas has zero entropy once conditioned on the side information stored in his memory—he knows all there is to know about the gas. Conceptually, the thermodynamic reason for this difference is that the demon is able to extract work from the gas, whereas the macroscopic observer is not. Indeed, the demon can exploit the side information stored in his memory to design a perfect trap door opening schedule which, when executed, concentrates all the particles on one side of the box. (This process can itself be thought of as c-not gates acting in the other direction.) With now all particles concentrated on one side of the box, the demon can extract work by replacing the separator by a piston and letting the gas expand isothermally. (Of course, the memory register is still littered with all the information about the gas; resetting the register costs work according to Landauer's principle, which is where the demon pays back his extracted work if he wishes to operate cyclically [2, 3].)

The above example shows that a fully general framework of thermodynamics should be universally applicable from the point of view of any observer, accounting for any level of knowledge one might possess about a system.

Our framework is well suited for describing different observers. Consider two observers, Alice and Bob, who have distinct degrees of knowledge about a system: For instance, a microscopic observer (Alice) might have access to individual position and momenta of all the particles of a gas, whilst a macroscopic observer (Bob) only has access to partial information given by macroscopic physical quantities such as temperature, pressure, volume, etc. More generally, we assume that the system's microscopic state space $\mathscr{H}_A$ (which Alice has access to) is transformed by a completely positive, trace-nonincreasing map $\mathcal{F}_{A\to B}^{\mathcal{A}\to\mathcal{B}}$ to a state space $\mathscr{H}_B$ which is used by Bob to describe the situation (Figure 4). For instance, if the microscopic system can be embedded into a bipartite system $\mathscr{H}_K \otimes \mathscr{H}_N$ which stores respectively the macroscopically available information (which Bob has access to) and the microscopic information (which only Alice has access to), then Bob's observations can then be related to Alice's simply by tracing out the $\mathscr{H}_N$ system.

Suppose that Alice observes some microscopic dynamics happening within $\mathscr{H}_A$, and that this evolution is Γ-sub-preserving with a particular operator $\Gamma_A^{\mathcal{A}}$. How does this evolution appear to Bob? It turns out that for Bob, these maps are also Γ-preserving maps, but relative to *his* Γ operator, which is simply given as $\Gamma_B^{\mathcal{B}} = \mathcal{F}_{A\to B}^{\mathcal{A}\to\mathcal{B}}(\Gamma_A^{\mathcal{A}})$, that is, by transforming Alice's Γ operator into Bob's picture. Conversely, a map which appears as $\Gamma^{\mathcal{B}}$-preserving to Bob, will be observed by Alice as being $\Gamma^{\mathcal{A}}$-preserving.

In order to give a precise meaning to the above statements, it is necessary to define precisely how a state described by Bob can be translated back to Alice's picture. Indeed, there can be several possible states for Alice which are compatible with Bob's state. We resort to the notion of *recovery* with the *Petz recovery*

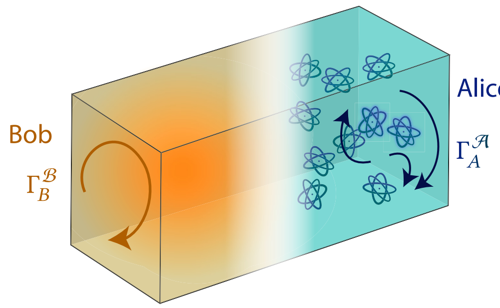

Figure 4. Observers in thermodynamics. Alice has access to microscopic degrees of freedom of a gas, while Bob can only observe its coarse macroscopic properties, such as its temperature T, volume V and pressure p. Alice describes the evolution of the gas using Gibbs-preserving maps, with a Gibbs state $\Gamma_A^{\mathcal{A}}$ on the full state space of the many particles of the gas. On the other hand, Bob describes the gas using his own knowledge—for instance, the macroscopic variables T, V, p—which in full generality we can represent as a quantum state in a state space $\mathscr{H}_B$ which is obtained by applying a given mapping $\mathcal{F}_{A\to B}^{\mathcal{A}\to\mathcal{B}}(\cdot)$ on Alice's state. (For instance, this map may trace out the inaccessible microscopic information.) States of the gas described by Bob may be transformed to Alice's picture by applying a Petz recovery map [58, 81–85]. Then, Alice's $\Gamma_A^{\mathcal{A}}$-preserving maps appear to Bob as $\Gamma_B^{\mathcal{B}}$-preserving maps, where Bob's $\Gamma_B^{\mathcal{B}}$ operator is taken to be $\Gamma_B^{\mathcal{B}} = \mathcal{F}_{A\to B}^{\mathcal{A}\to\mathcal{B}}(\Gamma_A^{\mathcal{A}})$. Conversely, operations which preserve $\Gamma_B^{\mathcal{B}}$ for Bob may be described by Alice as preserving $\Gamma_A^{\mathcal{A}}$.

map [58, 81–85]: this gives in a sense the "best guess" of what the state on $\mathscr{H}_A$ could be, given only knowledge of Bob's state on $\mathscr{H}_B$. The recovery map corresponding to $\mathcal{F}_{A\to B}^{\mathcal{A}\to\mathcal{B}}$ is given by

$$\mathcal{R}_{B\to A}^{\mathcal{B}\to\mathcal{A}}(\cdot) = \Gamma_A^{\mathcal{A}\,1/2}\, \mathcal{F}_{A\leftarrow B}^{\mathcal{A}\leftarrow\mathcal{B}\,\dagger}\big(\Gamma_B^{\mathcal{B}\,-1/2}\,(\cdot)\,\Gamma_B^{\mathcal{B}\,-1/2}\big)\,\Gamma_A^{\mathcal{A}\,1/2}\,, \quad (27)$$

where $\Gamma_B^{\mathcal{B}} = \mathcal{F}_{A\to B}^{\mathcal{A}\to\mathcal{B}}(\Gamma_A^{\mathcal{A}})$ and where $\mathcal{F}_{A\leftarrow B}^{\mathcal{A}\leftarrow\mathcal{B}\,\dagger}$ is the adjoint of the superoperator $\mathcal{F}_{A\to B}^{\mathcal{A}\to\mathcal{B}}$. The recovery map is completely positive, trace nonincreasing, and satisfies $\mathcal{R}_{B\to A}^{\mathcal{B}\to\mathcal{A}}(\Gamma_B^{\mathcal{B}}) \leqslant \Gamma_A^{\mathcal{A}}$.

Hence, given a trace-nonincreasing mapping $\mathcal{E}_A^{\mathcal{A}}$ in Alice's picture, we define Bob's description of the mapping as the composed map of transforming into Alice's picture, applying the map, and transforming back to Bob's picture:

$$\mathcal{E}_B^{\mathcal{B}} = \mathcal{F}_{A\to B}^{\mathcal{A}\to\mathcal{B}} \circ \mathcal{E}_A^{\mathcal{A}} \circ \mathcal{R}_{B\to A}^{\mathcal{B}\to\mathcal{A}}\,. \quad (28)$$

Our claim is the following: If $\mathcal{E}_A^{\mathcal{A}}$ satisfies $\mathcal{E}_A^{\mathcal{A}}(\Gamma_A^{\mathcal{A}}) \leqslant \Gamma_A^{\mathcal{A}}$, then $\mathcal{E}_B^{\mathcal{B}}$ satisfies $\mathcal{E}_B^{\mathcal{B}}(\Gamma_B^{\mathcal{B}}) \leqslant \Gamma_B^{\mathcal{B}}$. Conversely, if we are given a trace-nonincreasing mapping $\mathcal{E}_B^{\mathcal{B}}$ in Bob's picture, then this map is described in Alice's picture as the composed map of transforming to Bob's picture, applying the map, and transforming back:

$$\mathcal{E}_A^{\mathcal{A}} = \mathcal{R}_{B\to A}^{\mathcal{B}\to\mathcal{A}} \circ \mathcal{E}_B^{\mathcal{B}} \circ \mathcal{F}_{A\to B}^{\mathcal{A}\to\mathcal{B}}\,; \quad (29)$$

we assert that if $\mathcal{E}_B^{\mathcal{B}}(\Gamma_B^{\mathcal{B}}) \leqslant \Gamma_B^{\mathcal{B}}$, then $\mathcal{E}_A^{\mathcal{A}}(\Gamma_A^{\mathcal{A}}) \leqslant \Gamma_A^{\mathcal{A}}$.

The proof of both claims follows straightforwardly by inserting the above definitions while noting that $\mathcal{R}_{B\to A}^{\mathcal{B}\to\mathcal{A}}(\Gamma_B^{\mathcal{B}}) \leqslant \Gamma_A^{\mathcal{A}}$ and recalling that a completely positive map $\mathcal{E}$ is trace-nonincreasing if and only if $\mathcal{E}^\dagger(\mathbb{1}) \leqslant \mathbb{1}$.

A simple example is the relation of the microcanonical to the canonical ensemble, or that of thermal operations to noisy operations and the notion of Gibbs-rescaling [25, 28, 32]. More precisely, if Alice describes unitary dynamics within an energy eigenspace of the joint system and a large heat bath, then Bob describes the dynamics of the system alone as Gibbs-preserving maps. Consider a system S and a heat bath R, with respective Hamiltonians H_S and H_R and total Hamiltonian $H_{SR} = H_S + H_R$. Suppose that Alice has microscopic access to the heat bath, and hence describes the situation using the state space $A = S \otimes R$. Assume that the global state and evolution are constrained to unitaries within a subspace of fixed total energy E. This evolution is in particular Γ-sub-preserving if we choose $\Gamma_A^{\mathcal{A}} = P_{SR}^E$, where P_{SR}^E is the projector onto the eigenspace of H_{SR} corresponding to the energy E. On the other hand, Bob only has access to the system $B = S$. The mapping $\mathcal{F}^{\mathcal{A}\to\mathcal{B}}$, which relates Alice's point of view to Bob's, simply traces out the heat bath R. Bob then describes the operator $\Gamma_A^{\mathcal{A}}$ as

$$\Gamma_S^{\mathcal{B}} = \operatorname{tr}_R \Gamma_{SR}^{\mathcal{A}} = \sum_{E_S,k} g(E - E_S)\,|E_S,k\rangle\langle E_S,k|_S\,, \quad (30)$$

where $g(E_R)$ is the degeneracy of the energy eigenspace of the heat bath corresponding to the energy E_R, and where the vectors $\{|E_S,k\rangle_S\}$ are the energy eigenstates on S with a possible degeneracy index k. Following standard arguments in statistical mechanics, and as argued in Ref. [25], we have in typical situations and under mild assumptions $g(E - E_S) \propto e^{-\beta E_S}$, and we hence recover the standard canonical form in (30). In other words, Bob describes the dynamics on S as maps which preserve the Gibbs state.

The above reasoning can be seen as a rule for transforming one observer's picture into another; it remains important to analyze the situation in the picture which accurately describes the state of knowledge of the input state of the agent carrying out the operations. The pictures are equivalent when Alice's state of knowledge of A is no more than what B can recover using the recovery map, i.e., when her input state is exactly of the form $\mathcal{R}_{B\to A}^{\mathcal{B}\to\mathcal{A}}(\rho_B^{\mathcal{B}})$ where $\rho_B^{\mathcal{B}}$ is the state of the system in Bob's picture. However, not all actions that Alice can perform using $\Gamma_A^{\mathcal{A}}$-sub-preserving maps must induce a $\Gamma_B^{\mathcal{B}}$-sub-preserving effective map on B. Indeed, if Alice's input state is more refined, i.e., if she has more fine-grained information about the microscopic initial state than what Bob can infer, then her actions might appear to Bob as violating his idea of the second law of thermodynamics. In this case, Alice may indeed perform $\Gamma_A^{\mathcal{A}}$-sub-preserving operations which result in an effective mapping on B which is not $\Gamma_B^{\mathcal{B}}$-sub-preserving. Enter Maxwell's demon.

Our framework hence allows us to systematically analyze a variety of settings inspired by Maxwell's demon. Returning to our example depicted in Figure 3, we identify Alice as the demon and Bob as the macroscopic observer. Alice can perform Gibbs-preserving operations on the joint system of the gas system S and her memory register M, which for simplicity we choose to have a completely degenerate Hamiltonian $H_M = 0$ and thus $\Gamma_M = \mathbb{1}_M$. Bob, on the other hand, describes the gas alone using standard thermodynamic variables, say energy E, volume V, and number of particles N. To relate both points of view, we write the gas system S as a bipartite system $S = K \otimes N$ with states of the form $|E,V,N\rangle\langle E,V,N|_K \otimes \tau_N^{E,V,N}$, where $\tau_N^{E,V,N}$ is the microcanonical state corresponding to the macroscopic variables E, V, N. We have $\tau_N^{E,V,N} = P_N^{E,V,N}/\Omega(E,V,N)$, where $P_N^{E,V,N}$ projects onto the subspace of the microscopic system corresponding to the fixed E, V, N, and where the partition function is $\Omega(E,V,N) = \mathrm{tr}[P_N^{E,V,N}]$. Then, Bob's picture is obtained from Alice's by disregarding the memory register as well as the microscopic information, which corresponds to the mapping $\mathcal{F}^{\mathcal{A}\to\mathcal{B}}_{KNM\to K}(\cdot) = \mathrm{tr}_{MN}(\cdot)$. Alice uses the description $\Gamma^{\mathcal{A}}_{KNM} = \sum_{E,V,N} |E,V,N\rangle\langle E,V,N|_K \otimes P_N^{E,V,N} \otimes \mathbb{1}_M$ (see previous section). Bob, on the other hand, describes the gas using $\Gamma^{\mathcal{B}}_K = \mathcal{F}^{\mathcal{A}\to\mathcal{B}}_{KNM\to K}(\Gamma^{\mathcal{A}}_{KNM}) = d_M \sum \Omega(E,V,N)\, |E,V,N\rangle\langle E,V,N|_K$, where d_M is the dimension of the system M. Using the fact that $\mathcal{F}^{\mathcal{A}\leftarrow\mathcal{B}\,\dagger}_{KNM\leftarrow K}(\cdot) = (\cdot) \otimes \mathbb{1}_{NM}$, the recovery map corresponding to $\mathcal{F}^{\mathcal{A}\to\mathcal{B}}_{KNM\to K}$ is determined to be

$$\mathcal{R}^{\mathcal{B}\to\mathcal{A}}_{K\to KNM}(\cdot) = \left(R_{K\to KN}\left[(\cdot)\otimes \mathbb{1}_N\right] R^\dagger_{K\leftarrow KN}\right) \otimes \frac{\mathbb{1}_M}{d_M}\;, \tag{31}$$

where we have defined the operator

$$R_{K\to KN} = \sum_{E,V,N} |E,V,N\rangle\langle E,V,N|_K \otimes \frac{P_N^{E,V,N}}{\sqrt{\Omega(E,V,N)}}\;. \tag{32}$$

Importantly, the recovery map applied to any state of the form $|E,V,N\rangle_K$ gives

$$\begin{aligned}\mathcal{R}^{\mathcal{B}\to\mathcal{A}}_{K\to KNM}&(|E,V,N\rangle\langle E,V,N|_K)\\ &= |E,V,N\rangle\langle E,V,N|_K \otimes \tau_N^{E,V,N} \otimes \frac{\mathbb{1}_M}{d_M}\;,\end{aligned} \tag{33}$$

i.e., Bob assigns a standard thermal state to all systems he can't otherwise access. From Alice's perspective (the demon's), the memory register M starts in a pure state $|0\rangle_M$, in order to store the future results from observations of the gas. On the other hand, Bob has no way to infer this state from his macroscopic information. Because of this, Alice can design processes which are perfectly Γ-sub-preserving from her perspective, but which can trick Bob into thinking he is observing a violation of the second law (as described in Figure 3). Consider for concreteness the following procedure: Alice performs a unitary process mapping the state $|E,V,N\rangle\langle E,V,N|_K \otimes \tau_N^{E,V,N} \otimes |0\rangle\langle 0|_M$ to $|E,V/2,N\rangle\langle E,V/2,N|_K \otimes \tau_N^{E,V/2,N} \otimes (d_M^{-1}\mathbb{1}_M)$, where we assume that the system M has just the right dimension to store all the entropy resulting from mapping a state $\tau_N^{E,V,N}$ to the state $\tau_N^{E,V/2,N}$ of lower rank (we assume for simplicity that the rank of $\tau_N^{E,V/2,N}$ divides that of $\tau_N^{E,V,N}$, and thus $\Omega(E,V,N) = d_M\,\Omega(E,V/2,N)$). Alice's process is fully Γ preserving, because it is unitary and it commutes with $\Gamma^{\mathcal{A}}_{KNM}$. However, from Bob's perspective, the gas changed its state from $|E,V,N\rangle_K$ to $|E,V/2,N\rangle_K$, in a blatant violation of his idea of the second law of thermodynamics! Of course, a clever Bob would be led to infer that there exists some system (M) which has interacted with the gas and which has absorbed the surplus entropy. The point is, however, that Bob can still very well apply his laws of thermodynamics (in the form of the restriction imposed by Γ-sub-preserving maps) as long as Alice doesn't "actively mess with him." That is, any observer can consistently apply the laws of thermodynamics (in the form of our framework) from their perspective, using the restriction of Γ-sub-preserving maps for appropriately chosen Γ operators as long as this restriction indeed holds. A Γ-sub-preserving restriction inferred from coarse-graining a finer Γ-sub-preserving restriction fails exactly when the finer-grained observer actively makes use of their privileged microscopic access.

A further example illustrating the necessity of treating thermodynamics as an observer-dependent framework, and where our framework could be applied, is provided by Jaynes' beautiful treatment of the Gibbs Paradox [80].

IV. DISCUSSION

One could think that thermodynamics, as a physical theory by essence, would require physical concepts, such as energy or number of particles, to be built into the theory as is done in usual textbooks. Our results align with the opposite view, where thermodynamics is a generic framework itself agnostic of any physical quantities such as "energy," which can be applied to different physical situations, in the same spirit as previously proposed approaches [86–91]. The physical properties of the system, such as energy, temperature or number of particles, are accounted for in our framework only through the abstract Γ operator.

Our results reveal the following picture: In any situation where the evolution of the system happens to obey a restriction of the form of preserving a certain operator Γ, then our framework applies; purity may be invested to lift the restriction on any process, in the amount given by the coherent relative entropy; and depending on how Γ is defined, one may relate this abstract resource to a physical resource such as mechanical work. Furthermore, if the states of interest of our system form a class of states which happen to be reversibly interconvertible, the macroscopic laws of thermodynamics emerge, along with the relevant thermodynamic potential.

The core of the framework is the Γ-sub-preserving restriction imposed on the free operations. The Γ operator encodes all the relevant physics of the system considered. The restriction may come from any physical reason—for instance, by assuming that the evolution is modeled by thermal operations on the microscopic level, or by otherwise justifying or assuming that the spontaneous dynamics are thermalizing in an appropriate sense. Furthermore, Γ-sub-preservation may come about in any situation where one or several conserved physical quantities are being exchanged with a corresponding thermodynamic bath, in a natural generalization of thermal operations [36, 38, 39].

Our framework is not limited to usual thermodynamics: By considering the Γ operator as an abstract entity, all considerations in our framework are of a purely quantum information theoretic nature and make no explicit reference to any physical quantity. Information processing is an example: One can consider purity as a resource and impose that operations sub-preserve the identity operator; our framework applies by taking $\Gamma = \mathbb{1}$; in this way one can recover the max-entropy as the number of pure qubits required to perform data compression of a given state. Furthermore, our framework is technically convenient to work with. Being a Γ-sub-preserving map is a semidefinite constraint, and thus optimization problems over free operations may often be formulated as semidefinite programs which exhibit a rich structure and can be solved efficiently. We might hence expect that our framework is possibly also relevant in quantum information processing or other purely information-theoretic settings. For instance, we expect connections with single-shot notions of conditional mutual information [50, 92, 93], which in the i.i.d. case can also be expressed as a difference of quantum relative entropies. Our approach is also promising for calculating remainder terms in recovery of quantum information [58, 84, 94–97].

Although the goal of our paper is to derive a fundamental limitation on operations in quantum thermodynamics, one can also ask the question of whether this limit can be achieved within a physically well-motivated set of operations. Because our bound is given by an optimization over Gibbs-preserving maps, it is clear that there is one such map which will attain that bound (or get arbitrarily close). However, it is not clear under which conditions our bound can be approximately attained in a more practical or realistic regime such as thermal operations (possibly combined with additional resources), as is the case for a system described by a fully degenerate Hamiltonian [60] or for classical systems [98].

The question of achievability is related to coherence in the context of thermodynamic transformations, an issue of significant recent interest [42–47]. In particular, thermal operations do not allow the generation of a coherent superposition of energy levels, while this is allowed to some extent by Gibbs-preserving maps [98]. Our approach suggests a possible interpretation for why this is the case: with Γ-sub-preserving operations, one requires no assumption that the system in question is isolated—for instance, Γ could be the reduced state on one party of a joint Gibbs state of a strongly interacting bipartite system. Indeed, the example in Ref. [98] can be explained in this way [64, Section 4.4.4]. Hence, the question of whether Gibbs-preserving maps may be implemented approximately using a more practical framework such as thermal operations (perhaps under certain conditions), remains an open question. In a similar vein, one could study the effect of catalysis in our framework [44, 51, 99], although one would probably have to consider state transitions rather than logical processes. A closer study of this type of situation is expected to reveal connections with smoothed generalized free energies [100] and the notion of approximate majorization [101].

Finally, our framework can describe a system at any degree of coarse-graining, including intermediate scales between the microscopic and macroscopic regimes. We can consider for instance a small-scale classical memory element which stores information using many electrons or many spins (such as everyday hard drives): The electrons may need to be treated thermodynamically, but not the system as a whole, since we have control over the information-bearing degrees of freedom. Other such examples include Maxwell demon-type scenarios, which our frameworks allows to treat systematically. In other words, we provide a self-contained framework of thermodynamics which allows to make the dependence on the observer explicit, underscoring the fact that thermodynamics is a theory which is relative to the observer [80].

V. ACKNOWLEDGMENTS

We are grateful to Mario Berta, Fernando Brandão, Frédéric Dupuis, Lea Krämer Gabriel, and Jonathan Oppenheim for discussions. We acknowledge contributions from the Swiss National Science Foundation (SNSF) via the NCCR QSIT as well as project No. 200020_165843. PhF acknowledges support from the SNSF through Early PostDoc.Mobility Fellowship No. P2EZP2_165239 hosted by the Institute for Quantum Information and Matter (IQIM) at Caltech, as well as from the National Science Foundation.

APPENDIX

The appendices are structured as follows. Appendix A offers some preliminary definitions and notation conventions. In Appendix B we prove the properties of our framework outlined in the main text, namely that any trace-nonincreasing, Γ-sub-preserving map can be dilated to a trace-preserving, Γ-preserving map, as well as the equivalence of a class of battery models. Appendix C is dedicated to the definition and properties of the coherent relative entropy. Appendix D discusses the robustness of battery states to small perturbations. Finally, Appendix E provides a selection of miscellaneous technical tools which are used in the rest of the paper.

Appendix A: Technical Preliminaries

Let us first fix some notation. The state space of a quantum system S is a Hilbert space $\mathscr{H}_S$ (in this work, we deal exclusively with finite-dimensional spaces), the dimension of which we denote by $|S|$. A quantum state ρ_S of S is a positive semidefinite operator of unit trace acting on $\mathscr{H}_S$. A subnormalized quantum state ρ_S is defined as satisfying $\operatorname{tr}\rho_S \leqslant 1$. In this work, quantum states are normalized to unit trace unless otherwise stated. We use the notation $A \geqslant 0$ to indicate that an operator A is positive semidefinite, and $A \geqslant B$ to indicate that $A - B \geqslant 0$. For any positive semidefinite operator A_S acting on $\mathscr{H}_S$ corresponding to a system S, we denote by $\Pi_S^{A_S}$ the projector onto the support of A_S. Furthermore, all projectors considered in this work are Hermitian. For each system S with Hilbert space $\mathscr{H}_S$, we fix a basis which we denote by $\{|k\rangle_S\}$. This serves to define, between any two systems A and B of same dimension, a reference (not normalized) entangled ket $|\Phi\rangle_{A:B} := \sum_k |k\rangle_A \otimes |k\rangle_B$, as well as the partial transpose operation $t_{A\to B}(\cdot) = \mathrm{tr}_A[\Phi_{A:B}(\cdot)] = \sum_{kk'} \langle k|\cdot|k'\rangle_A\, |k'\rangle\langle k|_B$. Furthermore, for any operator $\Xi_A \geqslant 0$, a ket $|\Xi\rangle_{A:B}$ is a purification of Ξ_A if and only if there exists a ket $|\Phi^\Xi\rangle_{A:B}$ of the form $|\Phi^\Xi\rangle_{A:B} = \sum_j |\chi_j\rangle_A |\chi_j\rangle_B$ with orthonormal sets $\{|\chi_j\rangle_A\}, \{|\chi_j\rangle_B\}$ such that $|\Xi\rangle_{A:B} = \Xi_A^{1/2}|\Phi^\Xi\rangle_{A:B} = \Xi_B^{1/2}|\Phi^\Xi\rangle_{A:B}$ with $\Xi_A = \mathrm{tr}_B|\Xi\rangle\langle\Xi|_{A:B}$ and $\Xi_B = \mathrm{tr}_A|\Xi\rangle\langle\Xi|_{A:B}$ (Schmidt decomposition); the ket $|\Xi\rangle_{A:B}$ is normalized if and only if $\mathrm{tr}\,\Xi_A = 1$.

Throughout this paper, 'log' denotes the logarithm in base 2.

1. Logical process and process matrix

We denote by a *logical process* a full description of a logical mapping of input states to output states:

Logical process. *A* logical process *$\mathcal{E}_{X\to X'}$ is a completely positive, trace-preserving map, mapping Hermitian operators on $\mathscr{H}_X$ to Hermitian operators on $\mathscr{H}_{X'}$.*

A logical process along with an input state may be characterized by their *process matrix*, defined as the Choi-Jamiołkowski map of the completely positive map, weighted by the input state.

Process matrix. *Let $\mathcal{E}_{X\to X'}$ be a logical process, and let σ_X be a quantum state. Let R_X be a system described by a Hilbert space $\mathscr{H}_{R_X} \simeq \mathscr{H}_X$, and let $|\sigma\rangle_{XR_X} = \sigma_X^{1/2}\,|\Phi\rangle\langle\Phi|_{X:R_X}\,\sigma_X^{1/2}$ be a purification of σ_X. Then the* process matrix *corresponding to $\mathcal{E}_{X\to X'}$ and σ_X is defined as $\rho_{X'R_X} = \mathcal{E}_{X\to X'}(\sigma_{XR_X})$, where the identity process is understood on R_X. The process matrix is itself a normalized quantum state. The (unnormalized)* Choi matrix *of $\mathcal{E}_{X\to X'}$ is $E_{X'R_X} = \mathcal{E}_{X\to X'}(\Phi_{X:R_X})$, and satisfies $\mathrm{tr}_{X'}\, E_{X'R_X} \leqslant \mathbb{1}_{R_X}$.*

We further have the properties $\rho_{X'R_X} = \sigma_{R_X}^{1/2}\, E_{X'R_X}\, \sigma_{R_X}^{1/2}$ and $\sigma_{R_X} = \mathrm{tr}_X\, \sigma_{XR_X} = t_{X\to R_X}(\sigma_X)$. Furthermore $\rho_R = \mathrm{tr}_{X'}\, \rho_{X'R_X} = \sigma_{R_X}$.

The process matrix in return fully determines the channel $\mathcal{E}_{X\to X'}$ on the support of σ_X, allowing for a full characterization of the input state as well as the logical process on the support of the input.

2. Distance measures on states

For two quantum states ρ, σ, the trace distance is given by $D(\rho,\sigma) = \frac{1}{2}\|\rho-\sigma\|_1$, and their fidelity is defined as $F(\rho,\sigma) = \mathrm{tr}\big[(\rho^{1/2}\,\sigma\,\rho^{1/2})^{1/2}\big]$. From the fidelity one can define the *purified distance*[1] as $P(\rho,\sigma) = \sqrt{1-F^2(\rho,\sigma)}$ [63, 69, 70].

It will also prove convenient to work with subnormalized quantum states. Following Refs. [63, 69, 70], for any two subnormalized states ρ, σ, we define the *(generalized) trace distance* $D(\rho,\sigma) = \frac{1}{2}\|\rho-\sigma\|_1 + \frac{1}{2}|\mathrm{tr}\,\rho - \mathrm{tr}\,\sigma|$, the *(generalized) fidelity* $F(\rho,\sigma) = \mathrm{tr}\big[(\rho^{1/2}\,\sigma\,\rho^{1/2})^{1/2}\big] + \sqrt{(1-\mathrm{tr}\,\rho)(1-\mathrm{tr}\,\sigma)}$ and the *(generalized) purified distance* $P(\rho,\sigma) = \sqrt{1-F^2(\rho,\sigma)}$. For any two subnormalized states ρ, σ, we have the useful relation $D(\rho,\sigma) \leqslant P(\rho,\sigma) \leqslant \sqrt{2\,D(\rho,\sigma)}$.

3. Semidefinite programming

Semidefinite programming is a useful toolbox which brings a rich structure to a certain class of optimization problems. We follow the notation of Refs. [104, 105], where proofs to the statements given here may also be found.

Let A and B be Hermitian matrices, let $\Phi(\cdot)$ be a Hermiticity-preserving superoperator, and let $X \geqslant 0$ be the optimization variable, which is a Hermitian matrix constrained to the cone of positive semidefinite matrices. The prototypical semidefinite program is an optimization problem of the following form:[2]

$$\text{minimize}: \quad \mathrm{tr}(AX) \tag{A.1a}$$
$$\text{subject to}: \quad \Phi(X) \geqslant B\,. \tag{A.1b}$$

To any such problem corresponds another, related problem in terms of a different variable $Y \geqslant 0$:

$$\text{maximize}: \quad \mathrm{tr}(BY) \tag{A.2a}$$
$$\text{subject to}: \quad \Phi^\dagger(Y) \leqslant A\,. \tag{A.2b}$$

The first problem is called the *primal problem*, and the second, *dual problem*. Either problem is deemed *feasible* if there exists a valid choice of the optimization variable satisfying the corresponding constraint. If there exists a $X \geqslant 0$ such that $\Phi(X) - B$ is positive definite, the primal problem is said to be *strictly feasible*; the dual is *strictly feasible* if there is a $Y \geqslant 0$ such that $A - \Phi^\dagger(Y)$ is positive definite. For these two problems, we define their optimal attained values

$$\alpha = \inf\big\{\mathrm{tr}(AX) : \Phi(X) \geqslant B, X \geqslant 0\big\}\,; \tag{A.3a}$$
$$\beta = \sup\big\{\mathrm{tr}(BY) : \Phi^\dagger(Y) \leqslant A, Y \geqslant 0\big\}\,, \tag{A.3b}$$

with the convention that $\alpha = -\infty$ if the primal problem is not feasible and $\beta = +\infty$ if the dual problem is not feasible.

For any semidefinite program, we have $\alpha \geqslant \beta$, a property called *weak duality*. This convenient relation allows us to immediately bound the optimal attained value of one of the two problems by picking any valid candidate in the other.

For some pairs of problems, we may have $\alpha = \beta$. In those cases we speak of *strong duality*. This is often the case in practice. A useful result here is Slater's theorem, providing sufficient conditions for strong duality [104, Theorem 2.2].

Theorem 1 (Slater's conditions for strong duality)**.** *Consider any semidefinite program written in the form (A.1), and let its dual problem be given by (A.2). Then:*

[1] The purified distance is also called *Bures distance* (up to a factor of 2) [102] and coincides to second order with the quantum *angle* [103].

[2] Several equivalent prototypical forms for semidefinite programs exist in the literature.

(i) *if the primal problem is feasible and the dual is strictly feasible, then strong duality holds and there exists a valid choice X for the primal problem with* $\operatorname{tr}(AX) = \alpha$*;*

(ii) *if the dual problem is feasible and the primal is strictly feasible, then strong duality holds and there exists a valid choice Y for the dual problem with* $\operatorname{tr}(BY) = \beta$*.*

We note that strong duality in itself doesn't necessarily imply the existence of an optimal choice of variables attaining the infimum or supremum. The existence of optimal primal or dual choices may be explicitly stated by Slater's conditions, or may be deduced by an auxiliary argument such as if the constraints force the optimization region to be compact.

Appendix B: Properties of our framework

1. Dilation of Γ-sub-preserving maps to Γ-preserving maps

For two systems X, Y, and corresponding operators $\Gamma_X, \Gamma_Y \geqslant 0$, We say that a completely positive map $\Phi_{X\to Y}$ is *Γ-sub-preserving* if it satisfies $\Phi(\Gamma_X) \leqslant \Gamma_Y$. Similarly, $\Phi_{X\to Y}$ is *Γ-preserving* if it satisfies $\Phi(\Gamma_X) = \Gamma_Y$.

From a technical point of view, trace-preserving Γ-preserving maps don't handle nicely systems of varying sizes or with different Γ operators. For example, if X and Y are systems with $\operatorname{tr}\Gamma_X \neq \operatorname{tr}\Gamma_Y$, there may clearly be no Γ-preserving map from X to Y which is also trace preserving. It turns out that, by focusing on trace-nonincreasing Γ-sub-preserving maps instead, we may circumvent the issue in a physically justified way: A trace-nonincreasing Γ-sub-preserving map can always be seen as a restriction of a Γ-preserving map on a larger system. Furthermore, the ancillas we have to include in this dilation are prepared in, or finish up in, eigenstates of the respective Γ operators.

Proposition 2 (Dilation of Γ-sub-preserving maps)**.** *Let K and L be quantum systems with corresponding Γ_K and Γ_L. Let $\tilde{\Phi}_{K\to L}$ be a trace-nonincreasing, Γ-sub-preserving map. Choose two arbitrary eigenvectors $|\mathrm{k}\rangle_K$ and $|\mathrm{l}\rangle_L$ of Γ_K and Γ_L, respectively. Then there exists a qubit system $\mathscr{H}_Q$ with corresponding Γ_Q diagonal in a basis composed of two orthogonal states $\{|\mathrm{i}\rangle_Q, |\mathrm{f}\rangle_Q\}$, such that there exists a trace-preserving, Γ-preserving map $\Phi_{KLQ\to KLQ}$ satisfying*

$$\tilde{\Phi}_{K\to L}(\cdot) = \langle \mathrm{k\,f}|\ \Phi_{KLQ\to KLQ}\Big((\cdot)\otimes|\mathrm{l\,i}\rangle\langle\mathrm{l\,i}|_{LQ}\Big)\ |\mathrm{k\,f}\rangle_{KQ}\ . \tag{B.1}$$

Here, the joint Γ operator on K, L, Q is $\Gamma_{KLQ} = \Gamma_K \otimes \Gamma_L \otimes \Gamma_Q$. Furthermore, the corresponding eigenvalues satisfy

$$\langle \mathrm{l}|\Gamma_L|\mathrm{l}\rangle_L \langle \mathrm{i}|\Gamma_Q|\mathrm{i}\rangle_Q = \langle \mathrm{k}|\Gamma_K|\mathrm{k}\rangle_K \langle \mathrm{f}|\Gamma_Q|\mathrm{f}\rangle_Q\ . \tag{B.2}$$

(Proof on page 14.)

This means that for any trace-nonincreasing, Γ-sub-preserving map $\tilde{\Phi}_{K\to L}$, we may find a larger system and a trace-preserving, Γ-preserving map Φ_{KLQ} such that $\tilde{\Phi}_{K\to L}$ is seen as the restriction of Φ_{KLQ} to the case where the input is fixed to $|\mathrm{l\,i}\rangle_{LQ}$ on LQ and where the output is post-selected with $|\mathrm{k\,f}\rangle_{KQ}$ on KQ.

If the operators $\Gamma_K, \Gamma_L, \Gamma_Q$ come from Hamiltonians H_K, H_L, H_Q as $\Gamma_i = e^{-\beta H_i}$ for a fixed inverse temperature β, then the ancillas are prepared and left in pure energy eigenstates, specifically $|\mathrm{l\,i}\rangle_{LQ}$ for the input and $|\mathrm{k\,f}\rangle_{KQ}$ for the output. Furthermore condition (B.2) ensures that the total energy of the ancillas remains the same:

$$\langle \mathrm{l}|H_L|\mathrm{l}\rangle_L + \langle \mathrm{i}|H_Q|\mathrm{i}\rangle_Q = \langle \mathrm{k}|H_K|\mathrm{k}\rangle_K + \langle \mathrm{f}|H_Q|\mathrm{f}\rangle_Q\ . \tag{B.3}$$

2. Equivalence of battery models

Consider a logical process $\mathcal{E}_{X\to X'}$ which is not itself a free operation (i.e., $\mathcal{E}_{X\to X'}(\Gamma_X) \nleqslant \Gamma_{X'}$). It turns out that it is possible to implement this process by investing a certain amount of resources by means of an explicit battery system.

One example of such a battery system is the *information battery*. The information battery is a quantum system A of dimension which we denote by $|A|$, and for which $\Gamma_A = \mathbb{1}_A$. We require the battery to initially be prepared in a state $2^{-\lambda_1}\mathbb{1}_{2^{\lambda_1}}$ and to finish in a state $2^{-\lambda_2}\mathbb{1}_{2^{\lambda_2}}$ at the end, where both states are simply a state with a flat spectrum of rank 2^{λ_1} or 2^{λ_2}, and where we require that $\lambda_1, \lambda_2 \geqslant 0$ and that $2^{\lambda_1}, 2^{\lambda_2}$ are integers. If λ_1, λ_2 are themselves integers, this corresponds exactly to having λ_1 or λ_2 qubits in a fully mixed state and the remaining qubits in a pure state.

It is known that this model is equivalent to several other battery models known in the literature [28], notably the work bit (or "wit") [25, 28], or a "weight" system [26, 57]. Here, we point out that these models are in fact different instances of a more general description, making their equivalence manifest.

The most general system we have shown to be usable as a battery system is simply any system W with a arbitrary Γ_W operator, which is restricted to be in states of the form $\sigma = (P\Gamma_W P)/\operatorname{tr}P\Gamma_W$, where P is a projector which commutes with Γ_W. The "value" or "uselessness" of this state is given by the quantity $\log\operatorname{tr}(P\Gamma)$. The wit, the weight, as well as the information battery are all special cases of this general model.

The following proposition gives a necessary and sufficient condition as to when it is possible to lift the restriction given a particular charge state of the battery, and shows how the different battery systems are equivalent.

Proposition 3. *Let $\mathcal{T}_{X\to X'}$ be a completely positive, trace-nonincreasing map. Let $y \in \mathbb{R}$. Then, the following are equivalent:*

(i) *The map $\mathcal{T}_{X\to X'}$ satisfies*

$$\mathcal{T}_{X\to X'}(\Gamma_X) \leqslant 2^{-y}\,\Gamma_{X'}\ ; \tag{B.4}$$

(ii) *For any $\lambda_1, \lambda_2 \geqslant 0$ such that $2^{\lambda_1}, 2^{\lambda_2}$ are integers and $\lambda_1 - \lambda_2 \leqslant y$, there exists a large enough system A with $\Gamma_A = \mathbb{1}_A$ as well as a trace-nonincreasing, Γ-sub-preserving map $\Phi_{XA\to X'A}$ satisfying for all ω_X,*

$$\Phi_{XA\to X'A}\Big(\big(2^{-\lambda_1}\mathbb{1}_{2^{\lambda_1}}\big)\otimes\omega_X\Big) = \big(2^{-\lambda_2}\mathbb{1}_{2^{\lambda_2}}\big)\otimes\mathcal{T}_{X\to X'}(\omega_X)\ ; \tag{B.5}$$

(iii) *For a two-level system Q with two orthonormal states $|1\rangle_Q, |2\rangle_Q$, and with $\Gamma_Q = g_1|1\rangle\langle 1|_Q + g_2|2\rangle\langle 2|_Q$ chosen such that $g_2/g_1 \geqslant 2^{-y}$, there exists a trace-nonincreasing, Γ-sub-preserving map $\Phi'_{XQ\to X'Q}$ satisfying for all ω_X,*

$$\Phi'_{XQ\to X'Q}\Big(\omega_X \otimes |1\rangle\langle 1|_Q\Big) = \mathcal{T}_{X\to X'}(\omega_X) \otimes |2\rangle\langle 2|_Q\ ; \qquad \text{(B.6)}$$

(iv) *Let $\tilde{Q}$ be any system and choose two orthogonal states $|1\rangle_{\tilde{Q}}, |2\rangle_{\tilde{Q}}$ which are eigenstates of $\Gamma_{\tilde{Q}}$ corresponding to respective eigenvalues g_1, g_2 which satisfy $g_2/g_1 \geqslant 2^{-y}$. Then there exists a trace-nonincreasing, Γ-sub-preserving map $\Phi'_{X\tilde{Q}\to X'\tilde{Q}}$ satisfying for all ω_X,*

$$\Phi'_{X\tilde{Q}\to X'\tilde{Q}}\Big(\omega_X \otimes |1\rangle\langle 1|_{\tilde{Q}}\Big) = \mathcal{T}_{X\to X'}(\omega_X) \otimes |2\rangle\langle 2|_{\tilde{Q}}\ ; \qquad \text{(B.7)}$$

(v) *Let W_1, W_2 be quantum systems with respective corresponding Γ operators $\Gamma_{W_1}, \Gamma_{W_2}$, and let P_{W_1}, P'_{W_2} be projectors satisfying $[P_{W_1}, \Gamma_{W_1}] = 0$ and $[P'_{W_2}, \Gamma_{W_2}] = 0$, such that*

$$\frac{\operatorname{tr} P'_{W_2}\Gamma_{W_2}}{\operatorname{tr} P_{W_1}\Gamma_{W_1}} \geqslant 2^{-y}\ . \qquad \text{(B.8)}$$

Then there exists a Γ-preserving, trace-nonincreasing map $\Phi''_{XW_1\to X'W_2}$ such that for all ω_X,

$$\Phi''_{XW_1\to X'W_2}\left(\frac{P_{W_1}\Gamma_{W_1}P_{W_1}}{\operatorname{tr}(P_{W_1}\Gamma_{W_1})} \otimes \omega_X\right) = \frac{P'_{W_2}\Gamma_{W_2}P'_{W_2}}{\operatorname{tr}(P'_{W_2}\Gamma_{W_2})} \otimes \mathcal{T}_{X\to X'}(\omega_X)\ . \qquad \text{(B.9)}$$

(Proof on page 15.)

3. Proofs

Proof of Proposition 2. By definition, $\tilde{\Phi}_{K\to L}$ satisfies both $\tilde{\Phi}_{K\to L}(\Gamma_K) \leqslant \Gamma_L$ and $\tilde{\Phi}^\dagger_{K\leftarrow L}(\mathbb{1}_L) \leqslant \mathbb{1}_K$. Hence, let $F_K, G_L \geqslant 0$ such that

$$\tilde{\Phi}_{K\to L}(\Gamma_K) = \Gamma_L - G_L\ ; \qquad \text{(B.10a)}$$

$$\tilde{\Phi}^\dagger_{K\leftarrow L}(\mathbb{1}_L) = \mathbb{1}_K - F_K\ . \qquad \text{(B.10b)}$$

Let Π^Γ_L be the projector onto the support of Γ_L. We have $\Pi^\Gamma_L \leqslant \mathbb{1}_L$ and thus $\tilde{\Phi}^\dagger_{K\leftarrow L}\left(\Pi^\Gamma_L\right) \leqslant \tilde{\Phi}^\dagger_{K\leftarrow L}(\mathbb{1}_L) \leqslant \mathbb{1}_K$. So define $F'_K \geqslant 0$ such that

$$\tilde{\Phi}^\dagger_{K\leftarrow L}\left(\Pi^\Gamma_L\right) = \mathbb{1}_K - F'_K\ . \qquad \text{(B.10c)}$$

Let the system Q be as in the claim, with Γ_Q diagonal in the basis $\{|\mathrm{i}\rangle_Q, |\mathrm{f}\rangle_Q\}$. Define now the completely positive map

$$\begin{aligned}
&\Phi_{KLQ\to KLQ}(\cdot) = \qquad \text{(B.11)}\\
&\tilde{\Phi}_{K\to L}\Big(\langle \mathrm{l\,i}|\cdot|\mathrm{l\,i}\rangle_{LQ}\Big) \otimes |\mathrm{k\,f}\rangle\langle \mathrm{k\,f}|_{KQ}\\
&+ \Gamma_K^{1/2}\tilde{\Phi}^\dagger_{K\leftarrow L}\Big(\big(\Gamma_L^{-1/2}\langle \mathrm{k\,f}|_{KQ}\big)(\cdot)\big(\Gamma_L^{-1/2}|\mathrm{k\,f}\rangle_{KQ}\big)\Big)\Gamma_K^{1/2} \otimes |\mathrm{l\,i}\rangle\langle \mathrm{l\,i}|_{LQ}\\
&+ \Xi_{KL\to KL}\Big(\langle \mathrm{i}|\cdot|\mathrm{i}\rangle_Q\Big) \otimes |\mathrm{i}\rangle\langle \mathrm{i}|_Q\\
&+ \Omega_{KL\to KL}\Big(\langle \mathrm{f}|\cdot|\mathrm{f}\rangle_Q\Big) \otimes |\mathrm{f}\rangle\langle \mathrm{f}|_Q\ , \qquad \text{(B.12)}
\end{aligned}$$

with some completely positive maps $\Xi_{KL\to KL}$ and $\Omega_{KL\to KL}$ yet to be determined.

First, note that the property (B.1) is obvious for this Φ_{KLQ}, simply because $|\mathrm{i}\rangle_Q$ and $|\mathrm{f}\rangle_Q$ are orthogonal. It remains to exhibit explicit $\Xi_{KL\to KL}$ and $\Omega_{KL\to KL}$ such that Φ_{KLQ} is trace-preserving and Γ-preserving. Define as shorthands

$$\begin{aligned}
g_\mathrm{k} &= \langle \mathrm{k}|\Gamma_K|\mathrm{k}\rangle_K\ ; & g_\mathrm{l} &= \langle \mathrm{l}|\Gamma_L|\mathrm{l}\rangle_L\ ;\\
g_\mathrm{i} &= \langle \mathrm{i}|\Gamma_Q|\mathrm{i}\rangle_Q\ ; & g_\mathrm{f} &= \langle \mathrm{f}|\Gamma_Q|\mathrm{f}\rangle_Q\ .
\end{aligned} \qquad \text{(B.13)}$$

Note that Condition (B.2) is then equivalent to

$$g_\mathrm{l}\cdot g_\mathrm{i} = g_\mathrm{k}\cdot g_\mathrm{f}\ , \qquad \text{(B.14)}$$

and that this is straightforwardly satisfied for an appropriate choice of Γ_Q (and hence of g_i, g_f).

At this point, we'll derive conditions that $\Xi_{KL\to KL}$ and $\Omega_{KL\to KL}$ need to satisfy in order for $\Phi_{KLQ\to KLQ}$ to map Γ_{KLQ} onto itself and to be trace-preserving. Calculate

$$\begin{aligned}
&\Phi_{KLQ\to KLQ}(\Gamma_{KLQ})\\
&\quad= g_\mathrm{l} g_\mathrm{i}\,\tilde{\Phi}_{K\to L}(\Gamma_K) \otimes |\mathrm{k\,f}\rangle\langle \mathrm{k\,f}|_{KQ}\\
&\qquad+ g_\mathrm{k} g_\mathrm{f}\,\Gamma_K^{1/2}\tilde{\Phi}^\dagger_{K\leftarrow L}\left(\Pi^\Gamma_L\right)\Gamma_K^{1/2} \otimes |\mathrm{l\,i}\rangle\langle \mathrm{l\,i}|_{LQ}\\
&\qquad+ g_\mathrm{i}\Xi_{KL\to KL}(\Gamma_{KL}) \otimes |\mathrm{i}\rangle\langle \mathrm{i}|_Q + g_\mathrm{f}\Omega_{KL\to KL}(\Gamma_{KL}) \otimes |\mathrm{f}\rangle\langle \mathrm{f}|_Q\\
&\quad= |\mathrm{f}\rangle\langle \mathrm{f}|_Q \otimes \big[g_\mathrm{l} g_\mathrm{i}(\Gamma_L - G_L) \otimes |\mathrm{k}\rangle\langle \mathrm{k}|_K + g_\mathrm{f}\Omega_{KL\to KL}(\Gamma_{KL})\big]\\
&\qquad+ |\mathrm{i}\rangle\langle \mathrm{i}|_Q \otimes \Big[g_\mathrm{k} g_\mathrm{f}\Gamma_K^{1/2}(\mathbb{1}_K - F'_K)\Gamma_K^{1/2} \otimes |\mathrm{l}\rangle\langle \mathrm{l}|_{KQ}\\
&\qquad\qquad+ g_\mathrm{i}\Xi_{KL\to KL}(\Gamma_{KL})\Big]\ . \qquad \text{(B.15)}
\end{aligned}$$

We see that in order for this last expression to equal $\Gamma_{KLQ} = g_\mathrm{f}|\mathrm{f}\rangle\langle \mathrm{f}|_Q \otimes \Gamma_{KL} + g_\mathrm{i}|\mathrm{i}\rangle\langle \mathrm{i}|_Q \otimes \Gamma_{KL}$, we need that the terms in square brackets above obey

$$g_\mathrm{l} g_\mathrm{i}(\Gamma_L - G_L) \otimes |\mathrm{k}\rangle\langle \mathrm{k}|_K + g_\mathrm{f}\Omega_{KL\to KL}(\Gamma_{KL}) = g_\mathrm{f}\,\Gamma_{KL}\ ; \qquad \text{(B.16a)}$$

$$g_\mathrm{k} g_\mathrm{f}\Gamma_K^{1/2}(\mathbb{1}_K - F'_K)\Gamma_K^{1/2} \otimes |\mathrm{l}\rangle\langle \mathrm{l}|_{KQ} + g_\mathrm{i}\Xi_{KL\to KL}(\Gamma_{KL}) = g_\mathrm{i}\,\Gamma_{KL}\ . \qquad \text{(B.16b)}$$

On the other hand, the adjoint map of $\Phi_{KLQ\to KLQ}$ is relatively straightforward to identify as

$$\begin{aligned}
&\Phi^\dagger_{KLQ\leftarrow KLQ}(\cdot) =\\
&\tilde{\Phi}^\dagger_{K\leftarrow L}\Big(\langle \mathrm{k\,f}|\cdot|\mathrm{k\,f}\rangle_{KQ}\Big) \otimes |\mathrm{l\,i}\rangle\langle \mathrm{l\,i}|_{LQ}\\
&+ \Gamma_L^{-1/2}\tilde{\Phi}_{K\to L}\Big(\big(\Gamma_K^{1/2}\langle \mathrm{l\,i}|_{LQ}\big)(\cdot)\big(\Gamma_K^{1/2}|\mathrm{l\,i}\rangle_{LQ}\big)\Big)\Gamma_L^{-1/2} \otimes |\mathrm{k\,f}\rangle\langle \mathrm{k\,f}|_{KQ}\\
&+ \Xi^\dagger_{KL\leftarrow KL}\Big(\langle \mathrm{i}|\cdot|\mathrm{i}\rangle_Q\Big) \otimes |\mathrm{i}\rangle\langle \mathrm{i}|_Q\\
&+ \Omega^\dagger_{KL\leftarrow KL}\Big(\langle \mathrm{f}|\cdot|\mathrm{f}\rangle_Q\Big) \otimes |\mathrm{f}\rangle\langle \mathrm{f}|_Q\ . \qquad \text{(B.17)}
\end{aligned}$$

We may thus now derive the conditions on $\Xi_{KL\to KL}$ and $\Omega_{KL\to KL}$ for $\Phi_{KLQ\to KLQ}$ to be trace-preserving. Specifically, we need to ensure that $\Phi^\dagger_{KLQ\leftarrow KLQ}(\mathbb{1}_{KLQ}) = \mathbb{1}_{KLQ}$. A calculation gives us

$$\begin{aligned}
&\Phi^\dagger_{KLQ\leftarrow KLQ}(\mathbb{1}_{KLQ})\\
&\quad= \tilde{\Phi}^\dagger_{K\leftarrow L}(\mathbb{1}_L) \otimes |\mathrm{l\,i}\rangle\langle \mathrm{l\,i}|_{LQ}\\
&\qquad+ \Gamma_L^{-1/2}\tilde{\Phi}_{K\to L}(\Gamma_K)\Gamma_L^{-1/2} \otimes |\mathrm{k\,f}\rangle\langle \mathrm{k\,f}|_{KQ}\\
&\qquad+ \Xi^\dagger_{KL\leftarrow KL}(\mathbb{1}_{KL}) \otimes |\mathrm{i}\rangle\langle \mathrm{i}|_Q\\
&\qquad+ \Omega^\dagger_{KL\leftarrow KL}(\mathbb{1}_{KL}) \otimes |\mathrm{f}\rangle\langle \mathrm{f}|_Q\ .\\
&\quad= |\mathrm{f}\rangle\langle \mathrm{f}|_Q \otimes \Big[\Gamma_L^{-1/2}(\Gamma_L - G_L)\Gamma_L^{-1/2} \otimes |\mathrm{k}\rangle\langle \mathrm{k}|_K\\
&\qquad+ \Omega^\dagger_{KL\leftarrow KL}(\mathbb{1}_{KL})\Big]\\
&\qquad+ |\mathrm{i}\rangle\langle \mathrm{i}|_Q \otimes \Big[(\mathbb{1}_K - F_K) \otimes |\mathrm{l}\rangle\langle \mathrm{l}|_L\\
&\qquad+ \Xi^\dagger_{KL\leftarrow KL}(\mathbb{1}_{KL})\Big]\ . \qquad \text{(B.18)}
\end{aligned}$$

Thus, for $\Phi_{KLQ\to KLQ}$ to be trace-preserving we must have

$$\Gamma_L^{-1/2}\,(\Gamma_L - G_L)\,\Gamma_L^{-1/2} \otimes |\mathrm{k}\rangle\langle\mathrm{k}|_K + \Omega^\dagger_{KL\leftarrow KL}\,(\mathbb{1}_{KL}) = \mathbb{1}_{KL}\ ; \tag{B.19a}$$

$$(\mathbb{1}_K - F_K) \otimes |\mathrm{l}\rangle\langle\mathrm{l}|_L + \Xi^\dagger_{KL\leftarrow KL}\,(\mathbb{1}_{KL}) = \mathbb{1}_{KL}\ . \tag{B.19b}$$

Let us now explicitly construct an $\Xi_{KL\to KL}$ which satisfies both (B.16b) and (B.19b). These conditions may be written as

$$\Xi_{KL\to KL}\,(\Gamma_{KL}) = \Gamma_{KL} - g_\mathrm{l}\,\Gamma_K^{1/2}\,(\mathbb{1}_K - F'_K)\,\Gamma_K^{1/2} \otimes |\mathrm{l}\rangle\langle\mathrm{l}|_L =: A_{KL}\ ; \tag{B.20a}$$

$$\Xi^\dagger_{KL\leftarrow KL}\,(\mathbb{1}_{KL}) = \mathbb{1}_{KL} - (\mathbb{1}_K - F_K) \otimes |\mathrm{l}\rangle\langle\mathrm{l}|_L =: B_{KL} \tag{B.20b}$$

where we have used (B.14) and defined two new operators A_{KL} and B_{KL}. Observe now that since $g_\mathrm{l}\,\Gamma_K^{1/2}\left(\mathbb{1}_K - F'_K\right)\Gamma_K^{1/2} \otimes |\mathrm{l}\rangle\langle\mathrm{l}|_L \leqslant \Gamma_K \otimes (g_\mathrm{l}\,|\mathrm{l}\rangle\langle\mathrm{l}|_L) \leqslant \Gamma_{KL}$, we have that $A_{KL} \geqslant 0$. Similarly, $(\mathbb{1}_K - F_K) \otimes |\mathrm{l}\rangle\langle\mathrm{l}|_L \leqslant \mathbb{1}_{KL}$ and hence $B_{KL} \geqslant 0$. Let ξ_{KL} be a quantum state defined as follows: If $\mathrm{tr}\,A_{KL} \neq 0$, then $\xi_{KL} = A_{KL}/\mathrm{tr}\,A_{KL}$; else $\xi_{KL} = \mathbb{1}_{KL}/|KL|$. Then define

$$\Xi_{KL\to KL}\,(\cdot) = \mathrm{tr}\,(B_{KL}\ (\cdot))\ \xi_{KL}\ . \tag{B.21}$$

We then have

$$\Xi^\dagger_{KL\leftarrow KL}\,(\mathbb{1}_{KL}) = \mathrm{tr}\,(\xi_{KL}\,\mathbb{1}_{KL})\ B_{KL} = B_{KL}\ , \tag{B.22}$$

thus satisfying condition (B.20b). On the other hand we have

$$\Xi_{KL\to KL}\,(\Gamma_{KL}) = \mathrm{tr}\,[B_{KL}\,\Gamma_{KL}]\ \xi_{KL}\ , \tag{B.23}$$

which we need to show equals A_{KL} to satisfy condition (B.20a). Consider first the case where $\mathrm{tr}\,A_{KL} = 0$ and hence $A_{KL} = 0$. Then $\Gamma_{KL} = g_\mathrm{l}\Gamma_K^{1/2}\left(\mathbb{1}_K - F'_K\right)\Gamma_K^{1/2} \otimes |\mathrm{l}\rangle\langle\mathrm{l}|_L$, and hence $\Gamma_L = g_\mathrm{l}|\mathrm{l}\rangle\langle\mathrm{l}|_L$ and $F'_K = 0$. Since $\tilde\Phi^\dagger_{K\leftarrow L}(\Pi^\Gamma_L) \leqslant \tilde\Phi^\dagger_{K\leftarrow L}\,(\mathbb{1}_L)$, we have $F_K \leqslant F'_K$ and thus $F_K = 0$. Then $B_{KL} = \mathbb{1}_K \otimes (\mathbb{1}_L - |\mathrm{l}\rangle\langle\mathrm{l}|_L)$. Thus, B_{KL} has no overlap with $\Gamma_{KL} = \Gamma_K \otimes (g_\mathrm{l}|\mathrm{l}\rangle\langle\mathrm{l}|_L)$ and (B.23) $= 0 = A_{KL}$ as required. Now consider the case where $\mathrm{tr}\,A_{KL} \neq 0$. We have

$$\begin{aligned}\mathrm{tr}\,A_{KL} &= \mathrm{tr}\,\Gamma_{KL} - g_\mathrm{l}\,\mathrm{tr}\left[\left(\mathbb{1}_K - F'_K\right)\Gamma_K\right]\\ &= \mathrm{tr}\,\Gamma_{KL} - g_\mathrm{l}\,\mathrm{tr}\left[\tilde\Phi^\dagger_{K\leftarrow L}\left(\Pi^\Gamma_L\right)\Gamma_K\right]\\ &= \mathrm{tr}\,\Gamma_{KL} - g_\mathrm{l}\,\mathrm{tr}\left[\Pi^\Gamma_L\,\tilde\Phi_{K\to L}\,(\Gamma_K)\right]\ .\end{aligned} \tag{B.24}$$

Now, because $\tilde\Phi_{K\to L}\,(\Gamma_K) \leqslant \Gamma_L$, the operator $\tilde\Phi_{K\to L}\,(\Gamma_K)$ must lie within the support of Γ_L. Thus the projector in the last term of (B.24) has no effect and can be replaced by an identity operator. We then have

$$\begin{aligned}(\text{B.24}) &= \mathrm{tr}\,\Gamma_{KL} - g_\mathrm{l}\,\mathrm{tr}\left[\mathbb{1}_L\tilde\Phi_{K\to L}\,(\Gamma_K)\right]\\ &= \mathrm{tr}\,\Gamma_{KL} - g_\mathrm{l}\,\mathrm{tr}\left[\tilde\Phi^\dagger_{K\leftarrow L}\,(\mathbb{1}_L)\,\Gamma_K\right]\\ &= \mathrm{tr}\,\Gamma_{KL} - g_\mathrm{l}\,\mathrm{tr}\left[(\mathbb{1}_K - F_K)\,\Gamma_K\right]\\ &= \mathrm{tr}\,\Gamma_{KL} - \mathrm{tr}\left[(\mathbb{1}_K - F_K) \otimes |\mathrm{l}\rangle\langle\mathrm{l}|_L\Gamma_{KL}\right]\\ &= \mathrm{tr}\,(B_{KL}\Gamma_{KL})\ .\end{aligned} \tag{B.25}$$

Since $\mathrm{tr}(B_{KL}\Gamma_{KL}) = \mathrm{tr}(A_{KL})$, we have (B.23) $= A_{KL}$ as required. We have thus constructed $\Xi_{KL\to KL}$ such that it satisfies conditions (B.16b) and (B.19b).

Let's now proceed analogously for $\Omega_{KL\to KL}$. We can rewrite conditions (B.16a) and (B.19a) as

$$\Omega_{KL\to KL}\,(\Gamma_{KL}) = \Gamma_{KL} - g_\mathrm{k}|\mathrm{k}\rangle\langle\mathrm{k}|_K \otimes (\Gamma_L - G_L) =: C_{KL}\ ; \tag{B.26}$$

$$\Omega^\dagger_{KL\leftarrow KL}\,(\mathbb{1}_{KL}) = \mathbb{1}_{KL} - |\mathrm{k}\rangle\langle\mathrm{k}|_K \otimes \Gamma_L^{-1/2}\,(\Gamma_L - G_L)\,\Gamma_L^{-1/2} =: D_{KL}\ , \tag{B.27}$$

defining the operators C_{KL} and D_{KL}. We have $g_\mathrm{k}\,|\mathrm{k}\rangle\langle\mathrm{k}|_K \otimes (\Gamma_L - G_L) \leqslant \Gamma_{KL}$ and thus $C_{KL} \geqslant 0$. Also $\Gamma_L^{-1/2}\,(\Gamma_L - G_L)\,\Gamma_L^{-1/2} \leqslant \mathbb{1}_L$ and thus $D_{KL} \geqslant 0$. Proceeding as for $\Xi_{KL\to KL}$, let ω_{KL} be a quantum state defined as $\omega_{KL} = C_{KL}/\mathrm{tr}\,C_{KL}$ if $\mathrm{tr}\,C_{KL} \neq 0$ or $\omega_{KL} = \mathbb{1}_{KL}/|KL|$ otherwise. Define

$$\Omega_{KL\to KL}\,(\cdot) = \mathrm{tr}\,(D_{KL}\ (\cdot))\,\omega_{KL}\ . \tag{B.28}$$

Then

$$\Omega^\dagger_{KL\leftarrow KL}\,(\mathbb{1}_{KL}) = \mathrm{tr}\,(\omega_{KL}\mathbb{1}_{KL})\ D_{KL} = D_{KL}\ , \tag{B.29}$$

which satisfies (B.27). On the other hand, we have

$$\Omega_{KL\to KL}\,(\Gamma_{KL}) = \mathrm{tr}\,(D_{KL}\Gamma_{KL})\ \omega_{KL}\ , \tag{B.30}$$

which we need to show is equal to C_{KL}. First consider the case where $\mathrm{tr}\,C_{KL} = 0$, i.e. $C_{KL} = 0$. Then $\Gamma_{KL} = g_\mathrm{k}|\mathrm{k}\rangle\langle\mathrm{k}|_K \otimes (\Gamma_L - G_L)$, implying that $\Gamma_K = g_\mathrm{k}|\mathrm{k}\rangle\langle\mathrm{k}|_K$ and $G_L = 0$. Then $D_{KL} = \mathbb{1}_{KL} - |\mathrm{k}\rangle\langle\mathrm{k}|_K \otimes \Pi^{\Gamma_L}_L = \mathbb{1}_{KL} - \Pi^{\Gamma_{KL}}_{KL}$, and thus D_{KL} has no overlap with Γ_{KL}. It follows that (B.30) $= 0 = C_{KL}$ as required. Now assume that $\mathrm{tr}\,C_{KL} \neq 0$. Then

$$\begin{aligned}\mathrm{tr}\,(D_{KL}\Gamma_{KL}) &= \mathrm{tr}\,\Gamma_{KL} - g_\mathrm{k}\,\mathrm{tr}\left((\Gamma_L - G_L)\,\Pi^\Gamma_L\right)\\ &= \mathrm{tr}\,\Gamma_{KL} - g_\mathrm{k}\,\mathrm{tr}\,(\Gamma_L - G_L) = \mathrm{tr}\,C_{KL}\ ,\end{aligned} \tag{B.31}$$

where the projector Π^Γ_L has no effect in the second expression since $\Gamma_L - G_L$ is entirely contained within the support of Γ_L. Then again (B.30) $= C_{KL}$ as required.

We have thus constructed a completely positive, trace preserving map $\Phi_{KLQ\to KLQ}$ which maps Γ_{KLQ} onto itself and which satisfies (B.1). This concludes the proof. ∎

Proof of Proposition 3. The proof consists in showing (i) ⟹ (v) ⟹ (iv) ⟹ (iii) ⟹ (i) as well as (v) ⟹ (ii) ⟹ (i).

(i) ⟹ (v): By assumption we have $\mathcal{E}_{X\to X'}(\Gamma_X) \leqslant 2^{-y}\Gamma_{X'}$. Let Γ_{W_1}, Γ_{W_2} and P_{W_1}, P'_{W_2} satisfy the assumptions in the claim (v), and define the shorthands

$$\sigma^{(1)}_{W_1} = \frac{P_{W_1}\Gamma_{W_1}P_{W_1}}{\mathrm{tr}(P_{W_1}\Gamma_{W_1})}\ ; \qquad \sigma^{(2)}_{W_2} = \frac{P'_{W_2}\Gamma_{W_2}P'_{W_2}}{\mathrm{tr}(P'_{W_2}\Gamma_{W_2})}\ . \tag{B.32}$$

Define the map

$$\Phi''_{XW_1\to X'W_2}(\cdot) = \sigma^{(2)}_{W_2} \otimes \mathcal{E}_{X\to X'}\left[\mathrm{tr}_{W_1}\left(P_{W_1}(\cdot)\right)\right]\ . \tag{B.33}$$

This map is completely positive by construction, and is trace nonincreasing because it is a composition of trace nonincreasing maps. We need to show that it is Γ-sub-preserving. We have

$$\begin{aligned}\Phi''_{XW_1\to X'W_2}\left(\Gamma_X \otimes \Gamma_{W_1}\right) &= \left(\mathrm{tr}\,P_{W_1}\Gamma_{W_1}\right)\cdot\sigma^{(2)}_{W_2} \otimes \mathcal{E}_{X\to X'}(\Gamma_X)\\ &\leqslant 2^{-y}\,\frac{\mathrm{tr}\,P_{W_1}\Gamma_{W_1}}{\mathrm{tr}\,P'_{W_2}\Gamma_{W_2}}\cdot\left(P'_{W_2}\Gamma_{W_2}P'_{W_2}\right) \otimes \Gamma_{X'}\\ &\leqslant \Gamma_{W_2} \otimes \Gamma_{X'}\ ,\end{aligned} \tag{B.34}$$

using the fact that $P'_{W_2}\Gamma_{W_2}P'_{W_2} \leqslant \Gamma_{W_2}$ since Γ_{W_2} commutes with the projector P'_{W_2}.

(v) ⟹ (iv): This special case follows directly from (v) with $W_1 = W_2 = \bar{Q}$, $\Gamma_{W_1} = \Gamma_{W_2} = \Gamma_{\bar{Q}}$ and by choosing $P_{W_1} = |1\rangle\langle 1|_{\bar{Q}}$, $P'_{W_2} = |2\rangle\langle 2|_{\bar{Q}}$. Note that $g_1 = \mathrm{tr}\,P_{W_1}\Gamma_{W_1}$ and $g_2 = \mathrm{tr}\,P'_{W_2}\Gamma_{W_2}$ and hence indeed $\left(\mathrm{tr}\,P'_{W_2}\Gamma_{W_2}\right)/\left(\mathrm{tr}\,P_{W_1}\Gamma_{W_1}\right) = g_2/g_1 \geqslant 2^{-y}$.

(iv) ⟹ (iii): This is a trivial special case of (iv).

(iii) ⟹ (i): Let Γ_Q, $|1\rangle_Q$, $|2\rangle_Q$, g_1, g_2 and $\Phi'_{XQ\to X'Q}$ be any choices which satisfy the assumptions of (iii) and which also satisfy the choice $g_2/g_1 = 2^{-y}$. Observe that for any ω_X

$$\mathcal{E}_{X\to X'}(\omega_X) = \left\langle 2\middle|\Phi'_{XQ\to X'Q}\left(\omega_X \otimes |1\rangle\langle 1|_Q\right)\middle|2\right\rangle_Q\ . \tag{B.35}$$

Plugging in $\omega_X = \Gamma_X$, and using the fact that $g_1|1\rangle\langle 1|_Q \leqslant \Gamma_Q$ and that $\Phi'_{XQ\to X'Q}$ is Γ-sub-preserving,

$$\begin{aligned}\mathcal{E}_{X\to X'}(\Gamma_X) &\leqslant \left\langle 2\middle|g_1^{-1}\cdot\Phi'_{XQ\to X'Q}\left(\Gamma_X \otimes \Gamma_Q\right)\middle|2\right\rangle_Q\ .\\ &\leqslant \left\langle 2\middle|g_1^{-1}\cdot\Gamma_{X'} \otimes \Gamma_Q\middle|2\right\rangle_Q\ .\\ &= \frac{g_2}{g_1}\cdot\Gamma_{X'} = 2^{-y}\,\Gamma_{X'}\ .\end{aligned} \tag{B.36}$$

(v) $\Rightarrow$(ii): This is in fact another special case of (v). Let λ_1, λ_2 such that $\lambda_1 - \lambda_2 \leqslant y$ and that 2^{λ_1}, 2^{λ_2} are integers. Let A be any quantum system of dimension at least $\max\{2^{\lambda_1}, 2^{\lambda_2}\}$ and with $\Gamma_A = \mathbb{1}_A$. Now we use our assumption that (v) holds. Choose $W_1 = W_2 = A$, $P_{W_1} = \mathbb{1}_{2^{\lambda_1}}$, $P'_{W_2} = \mathbb{1}_{2^{\lambda_2}}$. Observe that $\operatorname{tr} P_{W_1}\Gamma_{W_1} = \operatorname{tr} P_{W_1} = 2^{\lambda_1}$ and $\operatorname{tr} P'_{W_2}\Gamma_{W_2} = \operatorname{tr} P'_{W_2} = 2^{\lambda_2}$, and hence the assumptions of the choices in (v) are satisfied. Then we know that there must exist a Γ-sub-preserving, trace-nonincreasing map $\Phi''_{XA\to X'A}$ obeying (B.9). The latter condition reads by plugging in our choices

$$\Phi''_{XA\to X'A}\big((2^{-\lambda_1}\mathbb{1}_{2^{\lambda_1}}) \otimes \omega_X\big) = \big(2^{-\lambda_2}\mathbb{1}_{2^{\lambda_2}}\big) \otimes \mathcal{E}_{X\to X'}(\omega_X) \tag{B.37}$$

for all ω_X. This is exactly the condition that Φ has to fulfill, and hence Φ may be taken equal to the map Φ''. It follows that (ii) is true.

(ii) $\Rightarrow$(i): Consider any $\lambda_1, \lambda_2 \geqslant 0$ with $\lambda_1 - \lambda_2 \leqslant y$. Let $\Phi_{XA\to X'A}$ be the corresponding Γ-sub-preserving map given by the assumption that (ii) holds. Observe that for all ω_X,

$$\mathcal{E}_{X\to X'}(\omega_X) = \operatorname{tr}_A\left\{\mathbb{1}_{2^{\lambda_2}}\ \Phi_{XA\to X'A}\big((2^{-\lambda_1}\mathbb{1}_{2^{\lambda_1}}) \otimes \omega_X\big)\right\}. \tag{B.38}$$

Plugging in $\omega_X = \Gamma_X$, and using the fact that Φ is Γ-sub-preserving,

$$\begin{aligned}\mathcal{E}_{X\to X'}(\Gamma_X) &\leqslant \operatorname{tr}_A\left\{\mathbb{1}_{2^{\lambda_2}}\ \Phi_{XA\to X'A}(2^{-\lambda_1}\cdot\Gamma_A\otimes\Gamma_X)\right\}\\ &\leqslant 2^{-\lambda_1}\cdot\operatorname{tr}_A\left\{\mathbb{1}_{2^{\lambda_2}}\ \Gamma_A\otimes\Gamma_{X'}\right\}\\ &= 2^{-(\lambda_1-\lambda_2)}\,\Gamma_{X'}\end{aligned} \tag{B.39}$$

The statement (i) then follows by choosing a sequence of (λ_1, λ_2) with $\lambda_1 - \lambda_2 \to y$. ∎

Appendix C: The coherent relative entropy

1. Definition and basic properties

Consider two quantum systems X and X', described by respective Γ operators Γ_X and $\Gamma_{X'}$. We would like to perform a logical process from X to X' which is described by the process matrix $\rho_{X'R_X}$. As we have seen, the process matrix uniquely identifies both an input state σ_X and a trace-nonincreasing, completely positive map $\mathcal{E}_{X\to X'}$ on the support of σ_X.

Because $\rho_{X'R_X}$ only fixes the mapping on the support of σ_X, there may be several trace-nonincreasing, completely positive maps $\mathcal{T}_{X\to X'}$ which implement this given process matrix. The coherent relative entropy is defined as the optimal battery usage achieved by a $\mathcal{T}_{X\to X'}$ with fixed process matrix $\rho_{X'R_X}$, relative to Γ operators $\Gamma_X, \Gamma_{X'}$.

In fact, we allow the implementation to fail with some fixed probability $\epsilon \geqslant 0$ which can be chosen freely. This allow us to ignore very improbable events. Such a practice is standard in the smooth entropy framework, and it is even necessary in order to make physical statements and recover the correct asymptotic behavior [61, 63, 69]. Hence, we allow the process matrix achieved by the optimization variable $\mathcal{T}_{X\to X'}$ on the given input state to only be ϵ-close to the requested process matrix $\rho_{X'R_X}$.

By Proposition 3, the optimal number of extracted battery charge y of a fixed $\mathcal{T}_{X\to X'}$ is given by the condition $\mathcal{T}_{X\to X'}(\Gamma_X) \leqslant 2^{-y}\,\Gamma_{X'}$. We are then directly led to the following definition.

Coherent Relative Entropy. *For a bipartite quantum normalized state $\rho_{X'R_X}$, two positive semidefinite operators Γ_X and $\Gamma_{X'}$ such that $t_{R_X\to X}(\rho_{X'R_X})$ lies in the support of $\Gamma_X\otimes\Gamma_{X'}$, and for $\epsilon \geqslant 0$, the* coherent relative entropy *is defined as*

$$\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) = \max_{\substack{\mathcal{T}_{X\to X'}(\Gamma_X)\leqslant 2^{-y}\,\Gamma_{X'}\\ \mathcal{T}^{\dagger}_{X\leftarrow X'}(\mathbb{1}_{X'})\leqslant\mathbb{1}_{R_X}\\ P(\mathcal{T}_{X\to X'}(\sigma_{XR_X}),\rho_{X'R_X})\leqslant\epsilon}} y\ , \tag{C.1}$$

where the optimization ranges over all $y \in \mathbb{R}$ and over all completely positive maps $\mathcal{T}_{X\to X'}$ satisfying the given conditions, and where we use the shorthand $|\sigma\rangle_{XR_X} = \rho^{1/2}_{R_X}\,|\Phi\rangle_{X:R_X}$.

If $\epsilon = 0$, we may omit the ϵ superscript altogether:

$$\hat{D}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) = \hat{D}^{\epsilon=0}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})\ . \tag{C.2}$$

Clearly, the coherent relative entropy is monotonously increasing in ϵ, as the optimization set gets larger.

We now introduce the variable $\alpha = 2^{-y}$ and denote by $T_{X'R_X}$ the Choi matrix of $\mathcal{T}_{X\to X'}$, allowing us to write the coherent relative entropy as a semidefinite program.

Proposition 4 (Explicit semidefinite program)**.** *For a bipartite quantum normalized state $\rho_{X'R_X}$, two positive semidefinite operators Γ_X and $\Gamma_{X'}$ such that $t_{R_X\to X}(\rho_{X'R_X})$ lies in the support of $\Gamma_X\otimes\Gamma_{X'}$, and for $\epsilon \geqslant 0$, the coherent relative entropy may be written as*

$$\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) = -\log\alpha\ , \tag{C.3}$$

where α is the optimal solution to the following semidefinite program in terms of the variables $T_{X'R_XE} \geqslant 0, \alpha \geqslant 0$, and dual variables $\mu, \omega_{X'}, X_{R_X} \geqslant 0$, with $|\rho\rangle_{X'R_XE}$ being an arbitrary but fixed purification of $\rho_{X'R_X}$ into an environment system E of dimension at least $|E| \geqslant |X'R_X|$:

Primal problem:

$$\begin{aligned}&\text{minimize:} && \alpha \\ &\text{subject to:} && \operatorname{tr}_{X'}\big[T_{X'R_X}\big] \leqslant \mathbb{1}_{R_X} && : X_{R_X} && \text{(C.4a)}\\ &&& \operatorname{tr}_{R_X}\big[T_{X'R_X}\,\Gamma_{R_X}\big] \leqslant \alpha\,\Gamma_{X'} && : \omega_{X'} && \text{(C.4b)}\\ &&& \operatorname{tr}(\rho^{1/2}_{R_X}\,T_{X'R_XE}\,\rho^{1/2}_{R_X}\,\rho_{X'R_XE}) \geqslant 1-\epsilon^2 && : Z_{X'R_X} && \text{(C.4c)}\end{aligned}$$

Dual problem:

$$\begin{aligned}&\text{maximize:} && \mu\,(1-\epsilon^2) - \operatorname{tr}(X_{R_X}) \\ &\text{subject to:} && \operatorname{tr}\left[\omega_{X'}\Gamma_{X'}\right] \leqslant 1 && : \alpha && \text{(C.5a)}\\ &&& \mu\,\rho^{1/2}_{R_X}\,\rho_{X'R_XE}\,\rho^{1/2}_{R_X} \leqslant \mathbb{1}_E\otimes\big(\omega_{X'}\otimes\Gamma_{R_X} + \mathbb{1}_{X'}\otimes X_{R_X}\big) && : T_{X'R_X} && \text{(C.5b)}\end{aligned}$$

where we use the shorthand $\Gamma_{R_X} = t_{X\to R_X}(\Gamma_X)$. (Proof on page 18.)

Note that the dual problem is strictly feasible (choose, e.g., $Z_{X'R_X} = 0$, $\omega_{X'} = 0$ and $X_{R_X} = \mathbb{1}_{R_X}$), and $T_{X'R_X} = \rho^{-1/2}_{R_X}\rho_{X'R_X}\rho^{-1/2}_{R_X}$ is a feasible primal candidiate, and hence by Slater's sufficiency conditions (Theorem 1) we have that strong duality holds and there always exists optimal primal candidates. For $\epsilon > 0$, the primal problem is also strictly feasible (choose $T_{X'R_X} = (1-\epsilon^2/2)\,\rho^{-1/2}_{R_X}\rho_{X'R_X}\rho^{-1/2}_{R_X}$), and there always exists optimal dual candidates as well. However, note that for $\epsilon = 0$ the primal problem is not always strictly feasible (indeed, constraint (C.4c) is very strong and fixes the mapping $T_{X'R_X}$ on a subspace; because it must be trace-preserving on that subspace then (C.4a) cannot be satisfied strictly). This means that there is a possibility that there is no choice of optimal dual variables. However, since strong duality holds, there is always a sequence of choices for dual variables whose attained objective value will converge to the optimal solution of the semidefinite program.

Here are first some basic properties of the coherent relative entropy.

Proposition 5 (Trivial bounds)**.** *For any $0 \leqslant \epsilon < 1$, we have*

$$\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \geqslant -\log \operatorname{tr} \Gamma_X - \log \|\Gamma_{X'}^{-1}\|_\infty - \log(1-\epsilon^2)\,; \quad \text{(C.6a)}$$

$$\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \leqslant \log \|\Gamma_X^{-1}\|_\infty + \log \operatorname{tr} \Gamma_{X'} - \log(1-\epsilon^2)\,. \quad \text{(C.6b)}$$

(Proof on page 18.)

In the thermodynamic version of the framework, these bounds can be understood in terms of work extraction. Suppose $\Gamma_X = \Gamma_{X'} = e^{-\beta H_X}$ with a Hamiltonian H_X and an inverse temperature β. Then $\log \|\Gamma_X^{-1}\|_\infty$ (resp. $\log \|\Gamma_{X'}^{-1}\|_\infty$) is β times the maximum energy of R (resp. X'), and similarly, $\operatorname{tr} \Gamma_X$ (resp. $\operatorname{tr} \Gamma_{X'}$) is the partition function of X (resp. X'). The partition function is directly related to the work cost of erasure (resp. formation) of a thermal state to (resp. from) a pure energy eigenstate of zero energy. In this case, the bounds (C.6) correspond to the ultimate worst and best cases respectively. The ultimate worst case is that we start off in a thermal state and end up in the highest energy level, whereas the absolute best case would be to start in the highest energy eigenstate and finish in the Gibbs state.

Much like the conditional entropy and relative entropy, the coherent relative entropy is invariant under partial isometries of which $\rho_{X'R}$ and Γ operators lie in the support. In particular, the coherent relative entropy is completely oblivious to dimensions of the Hilbert spaces which are not spanned by Γ_R and $\Gamma_{X'}$.

Proposition 6 (Invariance under isometries)**.** *Let $\tilde{X}$, $\tilde{X}'$ be new systems. Suppose there exist partial isometries $V_{X\to\tilde{X}}$ and $V'_{X'\to\tilde{X}'}$ such that both $t_{R_X\to X}(\rho_{R_X})$ and Γ_X are in the support of $V_{X\to\tilde{X}}$, and both $\rho_{X'}$ and $\Gamma_{X'}$ are in the support of $V'_{X'\to\tilde{X}'}$. Then*

$$\hat{D}^{\epsilon}_{\tilde{X}\to\tilde{X}'}((V'\otimes V)\,\rho_{X'R_X}\,(V'\otimes V)^\dagger \,\|\, V\Gamma_X V^\dagger, V'\Gamma_{X'}V'^\dagger) = \hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'})\,. \quad \text{(C.7)}$$

(Proof on page 19.)

This proposition allows us to embed states in larger dimensions, as well as to show that it is invariant under simultaneous action of unitaries on the states and the Γ operators.

We may also check the behavior of the coherent relative entropy under re-scaling of the Γ operators (as the latter need not conform to any normalization). Intuitively, in the thermodynamic case where $\Gamma = e^{-\beta H}$ for a Hamiltonian H and an inverse temperature β, the transformation $\Gamma \to a\Gamma$ for a constant factor a yields the Γ operator corresponding to the modified Hamiltonian $H \to H - \beta^{-1} \ln a$, that is, a constant energy shift of all levels. Consequently, we expect that scaling the Γ operators introduces a constant shift in the coherent relative entropy, which would correspond to providing the required energy to compensate for the global change in energy.

Proposition 7 (Scaling the Γ operators)**.** *For any $0 \leqslant \epsilon < 1$, and for real numbers $a, b > 0$,*

$$\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, a\Gamma_X, b\Gamma_{X'}) = \hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) + \log \frac{b}{a}\,. \quad \text{(C.8)}$$

(Proof on page 19.)

The coherent relative entropy furthermore obeys a superadditivity rule, expressing the fact that a joint implementation of two parallel independent processes cannot be worse than two separate implementations of each process.

Proposition 8 (Superadditivity for tensor products)**.** *Let systems X_1, X_1', X_2, X_2' have respective Γ operators Γ_{X_1}, $\Gamma_{X_1'}$, Γ_{X_2}, $\Gamma_{X_2'}$. Let $\rho_{X_1'R_{X_1}}$ and $\zeta_{X_2'R_{X_2}}$ be two quantum states. Then for any $\epsilon, \epsilon' \geqslant 0$,*

$$\hat{D}^{\epsilon''}_{X_1X_2\to X_1'X_2'}(\rho_{X_1'R_{X_1}} \otimes \zeta_{X_2'R_{X_2}} \,\|\, \Gamma_{X_1}\otimes\Gamma_{X_2}, \Gamma_{X_1'}\otimes\Gamma_{X_2'}) \geqslant \hat{D}^{\epsilon}_{X_1\to X_1'}(\rho_{X_1'R_{X_1}} \,\|\, \Gamma_{X_1}, \Gamma_{X_1'}) + \hat{D}^{\epsilon'}_{X_2\to X_2'}(\zeta_{X_2'R_{X_2}} \,\|\, \Gamma_{X_2}, \Gamma_{X_2'})\,, \quad \text{(C.9)}$$

where $\epsilon'' = \sqrt{\epsilon^2 + \epsilon'^2}$. *(Proof on page 19.)*

Perhaps surprisingly, we do not have equality in general in Proposition 8. One may see this with a simple example analogous to that in Ref. [98]. Consider two qubit systems Q_i with $\Gamma_{Q_i} = g_0|0\rangle\langle 0| + g_1|1\rangle\langle 1|$ (with $i = 1, 2$; $g_0 > g_1$). On a single system, performing the logical process $|0\rangle \to |+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ has a different cost than the yield of $|+\rangle \to |0\rangle$.[3] However, the transition $|0\rangle \otimes |+\rangle \to |+\rangle \otimes |0\rangle$ can be achieved with a swap operation, which is perfectly Γ-preserving and hence costs no pure qubits.

A further property of the coherent relative entropy can be derived in the case where the Γ operators are restricted by projecting them onto selected eigenkets, while still having the process matrix lying in their support. Then the coherent relative entropy remains unchanged.

Proposition 9 (Restricting the Γ operators)**.** *Let P_X and $P'_{X'}$ be projectors such that $[P_X, \Gamma_X] = 0$ and $[P'_{X'}, \Gamma_{X'}] = 0$. Define $\Gamma'_X = P_X \Gamma_X P_X$ and $\Gamma'_{X'} = P'_{X'}\Gamma_{X'}P'_{X'}$. Let $\rho_{X'R_X}$ be any quantum state with support inside that of $\Gamma'_{X'} \otimes \Gamma'_{R_X}$. Then*

$$\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma'_X, \Gamma'_{X'}) = \hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'})\,. \quad \text{(C.10)}$$

(Proof on page 19.)

[3] That the processes $|0\rangle \to |+\rangle$ and $|+\rangle \to |0\rangle$ have different work cost and yield respectively follows from Corollary 29 below. We have $D_{\mathrm{min},0}(|+\rangle\langle +| \,\|\, \Gamma) = -\log\langle +|\Gamma|+\rangle = -\log[(g_0+g_1)/2]$ and $-D_{\max}(|+\rangle\langle +| \,\|\, \Gamma) = -\log\|\Gamma^{-1/2}|+\rangle\langle +|\Gamma^{-1/2}\|_\infty = -\log\langle +|\Gamma^{-1}|+\rangle = -\log[(g_0^{-1}+g_1^{-1})/2]$ (the argument of the norm is a pure state).

Another property relates the coherent relative entropy to that with respect to different Γ operators which represent "at least or at most as much weight on each state," as represented as an operator inequality.

Proposition 10. *Let $\tilde{\Gamma}_X \geqslant 0$ and $\tilde{\Gamma}_{X'} \geqslant 0$ be such that $\tilde{\Gamma}_X \leqslant \Gamma_X$ and $\Gamma_{X'} \leqslant \tilde{\Gamma}_{X'}$. Then for any $\epsilon \geqslant 0$,*

$$\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \tilde{\Gamma}_X, \tilde{\Gamma}_{X'}) \geqslant \hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_R, \Gamma_{X'}) \;. \tag{C.11}$$

(Proof on page 19.)

We further note that it is possible to rewrite the definition of the coherent relative entropy in a slightly alternative form.

Proposition 11. *The optimization problem defining the coherent relative entropy can be rewritten as*

$$2^{-\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'})} = \min_{T_{X'R_X}} \left\| \Gamma_{X'}^{-1/2} \operatorname{tr}_{R_X}\left[T_{X'R_X}\, t_{X\to R_X}(\Gamma_X) \right] \Gamma_{X'}^{-1/2} \right\|_{\infty} \;, \tag{C.12}$$

where the minimization is taken over all positive semidefinite $T_{X'R_X}$ satisfying both conditions (C.4a) and (C.4c), and for which the operator $\operatorname{tr}_{R_X}\left(T_{X'R_X}\, t_{X\to R_X}(\Gamma_X)\right)$ lies within the support of $\Gamma_{X'}$. Equivalently,

$$2^{-\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'})} = \min_{\mathcal{T}_{X\to X'}} \left\| \Gamma_{X'}^{-1/2}\, \mathcal{T}_{X\to X'}[\Gamma_X]\, \Gamma_{X'}^{-1/2} \right\|_{\infty} \;, \tag{C.13}$$

where the minimization is taken over all trace nonincreasing, completely positive maps $\mathcal{T}_{X\to X'}$ which satisfy $P(\mathcal{T}_{X\to X'}[\sigma_{XR_X}], \rho_{X'R_X}) \leqslant \epsilon$ and for which $\mathcal{T}_{X\to X'}(\Gamma_X)$ lies within the support of $\Gamma_{X'}$. *(Proof on page 19.)*

Finally, we present an alternative form of the semidefinite program for the non-smooth coherent relative entropy, i.e., in the case where $\epsilon = 0$. This version of the semidefinite program will prove useful in some later proofs.

Proposition 12 (Non-smooth specialized semidefinite program)**.** *For a bipartite quantum state $\rho_{X'R_X}$, and two positive semidefinite operators Γ_X and $\Gamma_{X'}$ such that $t_{R_X\to X}(\rho_{X'R_X})$ lies in the support of $\Gamma_X \otimes \Gamma_{X'}$, the non-smooth coherent relative entropy can be written as*

$$\hat{D}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) = -\log \alpha \;; \tag{C.14}$$

where α is the optimal solution to the following semidefinite program in terms of the variables $T_{X'R_X} \geqslant 0, \alpha \geqslant 0$, and dual variables $Z_{X'R_X} = Z^{\dagger}_{X'R_X}, \omega_{X'} \geqslant 0, X_{R_X} \geqslant 0$:

Primal problem:

minimize: α

subject to:

$$\operatorname{tr}_{X'}\left[T_{X'R_X}\right] \leqslant \mathbb{1}_{R_X} \quad : X_{R_X} \tag{C.15a}$$

$$\operatorname{tr}_{R_X}\left[T_{X'R_X}\, t_{X\to R_X}(\Gamma_X)\right] \leqslant \alpha\, \Gamma_{X'} \quad : \omega_{X'} \tag{C.15b}$$

$$\rho_{R_X}^{1/2}\, T_{X'R_X}\, \rho_{R_X}^{1/2} = \rho_{X'R_X} \quad : Z_{X'R_X} \tag{C.15c}$$

Dual problem:

maximize: $\operatorname{tr}\left[Z_{X'R_X}\rho_{X'R_X}\right] - \operatorname{tr} X_{R_X}$

subject to:

$$\operatorname{tr}\left[\omega_{X'}\Gamma_{X'}\right] \leqslant 1 \quad : \alpha \tag{C.16a}$$

$$\rho_{R_X}^{1/2}\, Z_{X'R_X}\, \rho_{R_X}^{1/2} \leqslant t_{X\to R_X}(\Gamma_X) \otimes \omega_{X'} + X_{R_X} \otimes \mathbb{1}_{X'} \quad : T_{X'R_X} \tag{C.16b}$$

(Proof on page 19.)

Here are the proofs corresponding to this section's propositions.

Proof of Proposition 4. Write $|\sigma\rangle_{XR} = \rho_{R_X}^{1/2}|\Phi\rangle_{X:R_X}$. Let $|\rho\rangle_{X'R_XE}$ be any fixed purification of $\rho_{X'R_X}$ in an environment system E with dimension $|E| \geqslant |X'R_X|$.

First, consider any feasible candidates $T_{X'RE}$, α for (C.4). Then, setting $\mathcal{T}_{X\to X'}(\cdot) = \operatorname{tr}_E(T_{X'R_XE}\, t_{X\to R_X}(\cdot))$ and $y = -\log\alpha$ satisfies the requirements of (C.1), in particular, $F^2(\mathcal{T}_{X\to X'}(\sigma_{XR_X}), \rho_{X'R_X}) \geqslant \operatorname{tr}(\rho_{R_X}^{1/2}\, T_{X'R_XE}\, \rho_{R_X}^{1/2}\, \rho_{X'R_XE}) \geqslant 1-\epsilon^2$ by Uhlmann's theorem because $\rho_{R_X}^{1/2}\, T_{X'R_XE}\rho_{R_X}^{1/2}$ is a purification of $\mathcal{T}_{X\to X'}(\sigma_{XR_X})$.

Let $\mathcal{T}_{X\to X'}$ and y be valid candidates in (C.1). Thanks to Uhlmann's theorem, there exists a pure quantum state $|\tau\rangle_{X'R_XE}$ such that $F^2(\mathcal{T}_{X\to X'}(\sigma_{XR_X}), \rho_{X'R_X}) = \operatorname{tr}(\tau_{X'R_XE}\,\rho_{X'R_XE})$. Let $V_{X\to X'E}$ be a Stinespring dilation of $\mathcal{T}_{X\to X'}$, i.e., let $V_{X\to X'E}$ satisfy $V^{\dagger}V \leqslant \mathbb{1}_X$ and $\mathcal{T}_{X\to X'}(\cdot) = \operatorname{tr}_E(V_{X\to X'E}\,(\cdot)\,V^{\dagger})$. There exists a unitary W_E such that $|\tau\rangle_{X'R_XE} = W_E\, V_{X\to X'E}|\sigma\rangle_{X:R_X}$, since those two states are both purifications of $\mathcal{T}_{X\to X'}(\sigma_{XR_X})$. Now let $|T\rangle_{X'R_XE} = W_E\, V_{X\to X'E}|\Phi\rangle_{X:R_X}$ and $\alpha = 2^{-y}$. Then, $\operatorname{tr}_{X'E}(T_{X'R_XE}) = \operatorname{tr}_X(V^{\dagger}\, V\, \Phi_{X:R_X}) \leqslant \mathbb{1}_{R_X}$. Also, $\operatorname{tr}_{R_XE}[T_{X'R_XE}\,\Gamma_{R_X}] = \mathcal{T}_{X\to X'}(\Gamma_X) \leqslant 2^{-y}\,\Gamma_{X'} = \alpha\,\Gamma_{X'}$. Finally, $\operatorname{tr}(\rho_{R_X}^{1/2}\, T_{X'R_XE}\, \rho_{R_X}^{1/2}\, \rho_{X'R_XE}) = \operatorname{tr}(W_E\, V_{X\to X'E}\, \sigma_{XR}\, V^{\dagger} W_E^{\dagger}\, \rho_{X'R_XE}) = \operatorname{tr}(\tau_{X'R_XE}\,\rho_{X'R_XE}) = F^2(\mathcal{T}_{X\to X'}(\sigma_{XR_X}), \rho_{X'R_X}) \geqslant 1-\epsilon^2$. ∎

Proof of Proposition 5. Let $T_{X'R_XE} = (1-\epsilon^2)\,\rho_{R_X}^{-1/2}\rho_{X'R_XE}\rho_{R_X}^{-1/2}$ and note that the condition (C.4c) is fulfilled. On the other hand, $\operatorname{tr}_{X'E}T_{X'RE} = (1-\epsilon^2)\,\Pi_{R_X}^{\rho_{R_X}} \leqslant \mathbb{1}_{R_X}$ fulfilling (C.4a). Now observe that

$$\operatorname{tr}\left(T_{X'R_X}\,\Gamma_{R_X}\right) = (1-\epsilon^2)\operatorname{tr}\left(\Pi_{R_X}^{\rho_{R_X}}\,\Gamma_{R_X}\right) \leqslant (1-\epsilon^2)\operatorname{tr}\left(\Gamma_{R_X}\right) , \tag{C.17}$$

and hence $[(1-\epsilon^2)\operatorname{tr}(\Gamma_{R_X})]^{-1}\operatorname{tr}_{R_X}(T_{X'R_X}\,\Gamma_{R_X})$ is a subnormalized quantum state, which moreover lives within the support of $\Gamma_{X'}$ by assumption. Hence,

$$[(1-\epsilon^2)\operatorname{tr}(\Gamma_{R_X})]^{-1}\operatorname{tr}_{R_X}(T_{X'R_X}\,\Gamma_{R_X}) \leqslant \Pi_{X'}^{\Gamma_{X'}} \leqslant \|\Gamma_{X'}^{-1}\|_{\infty}\,\Gamma_{X'} , \tag{C.18}$$

noting that $\|\Gamma_{X'}^{-1}\|_{\infty}^{-1}$ is the minimal nonzero eigenvalue of $\Gamma_{X'}$. Thus, taking $\alpha = (1-\epsilon^2)\operatorname{tr}(\Gamma_{R_X})\,\|\Gamma_{X'}^{-1}\|_{\infty}$ satisfies (C.4b) yielding feasible primal candidates, which proves (C.6a).

Now consider the dual problem. Choosing $\omega_{X'} = (\operatorname{tr}\Gamma_{X'})^{-1}\mathbb{1}_{X'}$ immediately satisfies (C.5a). Using $\rho_{X'R_XE} \leqslant \mathbb{1}_{X'R_XE}$ and $\rho_{R_X} \leqslant \Pi_{R_X}^{\Gamma_{R_X}}$, we have

$$\begin{aligned} \mu\,\rho_{R_X}^{1/2}\rho_{X'R_XE}\rho_{R_X}^{1/2} &\leqslant \mu\,\Pi_{R_X}^{\Gamma_{R_X}} \otimes \mathbb{1}_{X'E} \\ &= \mu\,(\operatorname{tr}\Gamma_{X'})\,\mathbb{1}_E \otimes \omega_{X'} \otimes \Pi_R^{\Gamma_{R_X}} \\ &\leqslant \mu\,(\operatorname{tr}\Gamma_{X'})\,\|\Gamma_{R_X}^{-1}\|_{\infty}\,\mathbb{1}_E \otimes \omega_{X'} \otimes \Gamma_{R_X} , \end{aligned} \tag{C.19}$$

so we choose $\mu = (\operatorname{tr}\Gamma_{X'})^{-1}\,\|\Gamma_{R_X}^{-1}\|_\infty^{-1}$ and $X_{R_X} = 0$ in order to fulfill (c.5b), which proves (c.6b). ∎

Proof of Proposition 6. This is clearly the case, because the semidefinite problem lies entirely within the support of the isometries. Formally, any choice of variables for the original problem can be mapped in the new spaces through these partial isometries, and vice versa, and the attained values remain the same. Hence the optimal value of the problem is also the same. ∎

Proof of Proposition 7. Consider the optimal primal candidiates $T_{X'R_XE}$ and α for the problem defining $2^{-\hat{D}^\epsilon_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}$. Then $T_{X'R_XE}$ and $ab^{-1}\alpha$ are feasible primal candidates for the semidefinite program with the scaled Γ operators. Hence

$$2^{-\hat{D}^\epsilon_{X\to X'}(\rho_{X'R_X}\,\|\,a\Gamma_X,b\Gamma_{X'})} \leqslant \frac{a}{b}\alpha = \frac{a}{b}\,2^{-\hat{D}^\epsilon_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}\,. \tag{c.20}$$

The opposite direction follows by applying the same argument to the reverse situation with $\Gamma_X \to a^{-1}\Gamma_X, \Gamma_{X'} \to b^{-1}\Gamma_{X'}$. ∎

Proof of Proposition 8. Let $T_{X_1'R_{X_1}E_1}$, α_1 and $T_{X_2'R_{X_2}E_2}$, α_2 be the optimal choice of primal variables for $2^{-\hat{D}^\epsilon_{X_1\to X_1'}(\rho_{X_1'R_{X_1}}\,\|\,\Gamma_{X_1},\Gamma_{X_1'})}$ and $2^{-\hat{D}^{\epsilon'}_{X_2\to X_2'}(\zeta_{X_2'R_{X_2}}\,\|\,\Gamma_{X_2},\Gamma_{X_2'})}$, respectively. Now, let $\bar{T}_{X_1'X_2'R_{X_1}R_{X_2}E_1E_2} = T_{X_1'R_{X_1}E_1}\otimes T_{X_2'R_{X_2}E_2}$ and $\bar\alpha = \alpha_1\alpha_2$. Then

$$\operatorname{tr}_{R_{X_1}R_{X_2}}\left[\bar{T}_{X_1'X_2'R_{X_1}R_{X_2}}\Gamma_{R_{X_1}}\otimes\Gamma_{R_{X_2}}\right] \leqslant \alpha_1\alpha_2\,\Gamma_{X_1'}\otimes\Gamma_{X_2'}\,; \tag{c.21}$$

$$\operatorname{tr}_{X_1'X_2'}\left[\bar{T}_{X_1'X_2'R_{X_1}R_{X_2}}\right] \leqslant \mathbb{1}_{R_{X_1}}\otimes\mathbb{1}_{R_{X_2}}\,, \tag{c.22}$$

and

$$\begin{aligned}\operatorname{tr}\Big[(\rho^{1/2}_{R_{X_1}}\otimes\zeta^{1/2}_{R_{X_2}})\bar{T}_{X_1'X_2'R_{X_1}R_{X_2}E_1E_2}(\rho^{1/2}_{R_{X_1}}\otimes\zeta^{1/2}_{R_{X_2}})\,\cdot\\ \rho_{X_1'R_{X_1}E_1}\otimes\zeta_{X_2'R_{X_2}E_2}\Big] \geqslant (1-\epsilon^2)(1-\epsilon'^2) \geqslant 1-\epsilon''^2\,,\end{aligned} \tag{c.23}$$

and hence this choice of variables is feasible for the tensor product problem. We then have

$$\begin{aligned}&2^{-\hat{D}^{\epsilon''}_{X_1X_2\to X_1'X_2'}(\rho_{X_1'R_{X_1}}\otimes\zeta_{X_2'R_{X_2}}\,\|\,\Gamma_{X_1}\otimes\Gamma_{X_2},\Gamma_{X_1'}\otimes\Gamma_{X_2'})} \leqslant \alpha_1\alpha_2\\ &\quad= 2^{-\left[\hat{D}^\epsilon_{X_1\to X_1'}(\rho_{X_1'R_{X_1}}\,\|\,\Gamma_{X_1},\Gamma_{X_1'})+\hat{D}^{\epsilon'}_{X_2\to X_2'}(\zeta_{X_2'R_{X_2}}\,\|\,\Gamma_{X_2},\Gamma_{X_2'})\right]}\,.\end{aligned}$$ ∎

Proof of Proposition 9. Let $T_{X'R_XE}$ and α be the optimal feasible candidates for the primal semidefinite problem defining $2^{-\hat{D}^\epsilon_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}$. Let $T'_{X'R_XE} = (P'_{X'}\otimes P_{R_X})\,T_{X'RE}\,(P'_{X'}\otimes P_{R_X})$ and $\alpha' = \alpha$, writing $P_{R_X} = t_{X\to R_X}(P_X)$. Then

$$\begin{aligned}\operatorname{tr}_{X'}T'_{X'R_X} = P_{R_X}\operatorname{tr}_{X'}\left[P'_{X'}T_{X'R_X}\right]P_{R_X} &\leqslant P_{R_X}\operatorname{tr}_{X'}\left(T_{X'R_X}\right)P_{R_X}\\ &\leqslant P_{R_X} \leqslant \mathbb{1}_{R_X}\,,\end{aligned} \tag{c.24}$$

satisfying (c.4a), and

$$\begin{aligned}&\operatorname{tr}[\rho^{1/2}_{R_X}T'_{X'R_XE}\rho^{1/2}_{R_X}\,\rho_{X'R_XE}]\\ &\qquad= \operatorname{tr}[\rho^{1/2}_{R_X}T_{X'R_XE}\rho^{1/2}_{R_X}\,\rho_{X'R_XE}] \geqslant 1-\epsilon^2\,,\end{aligned} \tag{c.25}$$

where the first equality holds because ρ_{R_X} and $\rho_{X'R_XE}$ already lie within the support of P_{R_X} and $P'_{X'}\otimes P_{R_X}\otimes\mathbb{1}_E$, respectively, and hence those projectors have no effect. Hence (c.4c) is fulfilled. Now we have

$$\begin{aligned}\operatorname{tr}_{R_X}[T'_{X'R_X}\Gamma'_{R_X}] &= \operatorname{tr}_{R_X}[(P'_{X'}\otimes P_{R_X})T_{X'R_X}(P'_{X'}\otimes P_{R_X})\Gamma_{R_X}]\\ &\leqslant P'_{X'}\operatorname{tr}_{R_X}[T_{X'R_X}\Gamma_{R_X}]P'_{X'}\\ &\leqslant P'_{X'}(\alpha\Gamma_{X'})P'_{X'} = \alpha'\Gamma'_{X'}\,,\end{aligned} \tag{c.26}$$

using the fact that $\Gamma'_{R_X} \leqslant \Gamma_{R_X}$ (because $[P_{R_X},\Gamma_{R_X}] = 0$). Hence

$$2^{-\hat{D}^\epsilon_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma'_X,\Gamma'_{X'})} \leqslant 2^{-\hat{D}^\epsilon_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}\,. \tag{c.27}$$

Let μ, X_{R_X} and $\omega_{X'}$ be any dual feasible candidates for $2^{-\hat{D}^\epsilon_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}$. Now let $\mu' = \mu$, $X'_{R_X} = P_{R_X}X_{R_X}P_{R_X}$ and $\omega_{X'} = P'_{X'}\,\omega'_{X'}\,P'_{X'}$. Then $\operatorname{tr}(\omega'_{X'}\Gamma'_{X'}) = \operatorname{tr}(\omega_{X'}\Gamma'_{X'}) \leqslant \operatorname{tr}(\omega_{X'}\Gamma_{X'}) \leqslant 1$ (using the fact that $\Gamma'_{X'} \leqslant \Gamma_{X'}$ since $[\Gamma_{X'}, P'_{X'}] = 0$), in accordance with (c.5a). Also, apply $(P'_{X'}\otimes P_{R_X})(\cdot)(P'_{X'}\otimes P_{R_X})$ onto the dual constraint (c.5b) to immediately see that μ', $\omega'_{X'}$ and X_{R_X} obey the new constraint with Γ'_{R_X}. Finally, the attained dual value is

$$\mu'\,(1-\epsilon^2) - \operatorname{tr}(X'_{R_X}) \geqslant \mu\,(1-\epsilon^2) - \operatorname{tr}(X_{R_X})\,. \tag{c.28}$$

Hence, we now have

$$2^{-\hat{D}^\epsilon_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma'_X,\Gamma'_{X'})} \geqslant 2^{-\hat{D}^\epsilon_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}\,, \tag{c.29}$$

which completes the proof. ∎

Proof of Proposition 10. Let $T_{X'R_XE}$ and α be the optimal solution to the semidefinite program for $2^{-\hat{D}^\epsilon_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}$. They are then also feasible candidates for the semidefinite program for $2^{-\hat{D}^\epsilon_{X\to X'}(\rho_{X'R_X}\,\|\,\bar\Gamma_X,\bar\Gamma_{X'})}$, because the only condition that changes is (c.4b), which is obviously still satisfied. ∎

Proof of Proposition 11. Let $T_{X'R_X}$ be any candidate in the primal problem. If $\operatorname{tr}_R\big(T_{X'R_X}\big)$ does not lie within the support of $\Gamma_{X'}$, then condition (c.4b) is not satisfied and the candidate is not primal feasible; we can hence ignore it in the minimization. Otherwise, by conjugating condition (c.4b) by $\Gamma_{X'}^{-1/2}$, we see that (c.4b) is equivalent to

$$\Gamma_{X'}^{-1/2}\operatorname{tr}_{R_X}\left[T_{X'R_X}\,t_{X\to R_X}(\Gamma_X)\right]\Gamma_{X'}^{-1/2} \leqslant \alpha\,\Pi^{\Gamma_{X'}}_{X'}\,, \tag{c.30}$$

which in turn is equivalent to

$$\Gamma_{X'}^{-1/2}\operatorname{tr}_{R_X}\left[T_{X'R_X}\,t_{X\to R_X}(\Gamma_R)\right]\Gamma_{X'}^{-1/2} \leqslant \alpha\,\mathbb{1}\,, \tag{c.31}$$

because the left hand side of (c.30) is entirely within the support of its right hand side. Now, the optimal α which corresponds to this fixed $T_{X'R_X}$ is given by $\|\Gamma_{X'}^{-1/2}\operatorname{tr}_{R_X}\left[T_{X'R_X}\,t_{X\to R_X}(\Gamma_X)\right]\Gamma_{X'}^{-1/2}\|_\infty$. This chain of equivalences may be followed in reverse order, establishing the equivalence of the minimization problems.

The formulation in terms of channels follows immediately from the translation of one formalism to the other. ∎

Proof of Proposition 12. In the case $\epsilon = 0$, the conditions in (c.1) reduce to

$$\begin{aligned}\mathcal{T}_{X\to X'}(\Gamma_X) &\leqslant 2^{-y}\Gamma_{X'}\,;\\ \mathcal{T}^\dagger_{X\leftarrow X'}(\mathbb{1}_{X'}) &\leqslant \mathbb{1}_X\,;\\ \mathcal{T}_{X\to X'}(\sigma_{XR_X}) &= \rho_{X'R_X}\,,\end{aligned}$$

where we write $|\sigma\rangle_{XR_X} = \rho^{1/2}_{R_X}\,|\Phi\rangle_{X:R_X}$. These conditions, when written in terms of the Choi matrix $T_{X'R_X}$ corresponding to $\mathcal{T}_{X\to X'}$, yield precisely the semidefinite program given in the claim. ∎

2. Some special cases

In this section, we look at some instructive special cases where the coherent relative entropy can be evaluated exactly.

The first proposition concerns identity mappings. It is a property that one would expect very naturally: If the process matrix corresponds to the identity mapping on the support of the input, and if the Γ operators coincide, then the process should be a free operation and should not require a battery. This property may seem like a triviality, but it is in fact not so obvious to prove: Indeed, because the coherent relative entropy is a function of the process matrix only, the implementation can choose to implement whatever process it likes on the complement of the support of the input state. In other words, this proposition tells us that there is no way to extract work by exploiting the freedom on this complementary subspace when performing the identity map on the support of σ_X.

Proposition 13 (Identity mapping)**.** *Let $\mathrm{id}_{X\to X'}$ be the identity map from a system X to a system $X' \simeq X$. Assume that $\Gamma_{X'} = \mathrm{id}_{X\to X'}(\Gamma_X)$. Let σ_X be any state on X, let $R_X \simeq X$ and $|\sigma\rangle_{XR_X} = \sigma_X^{1/2}\,|\Phi\rangle_{X:R_X}$, and let $|\rho\rangle_{X'R_X}$ be the process matrix of the identity process applied on σ_X, i.e. $\rho_{X'R_X} = \mathrm{id}_{X\to X'}(\sigma_{XR_X})$. Then*

$$\hat{D}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) = 0\ . \tag{C.32}$$

Proof of Proposition 13. Let $\Phi_{X'R_X} = \mathrm{id}_{X\to X'}(\Phi_{X:R_X})$ be the unnormalized maximally entangled state on X' and R_X such that $\rho_{X'R_X} = \rho_{R_X}^{1/2}\Phi_{X'R_X}\rho_{R_X}^{1/2}$.

First we show that $\hat{D}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) \geqslant 0$. Consider the mapping $\mathcal{T}_{X\to X'} = \mathrm{id}_{X\to X'}$ and $y = 0$, i.e., consider the identity mapping as an implementation candidate. This clearly satisfies the requirements of the maximization in (C.1) for $\epsilon = 0$, and thus

$$\hat{D}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) \geqslant 0\ . \tag{C.33}$$

We prove the reverse direction by exhibiting dual candidates for the problem given in Proposition 12. The tricky part is that there might not be an optimal choice of dual variables. The best we can do in general is to come up with a sequence of choices for dual candidates whose attained value converges to 1. For any $\mu > 0$, let

$$Z_{X'R_X} = \mu\rho_{R_X}^{-1/2}\Phi_{X'R_X}\rho_{R_X}^{-1/2}\ ;\qquad \omega_{X'} = \left(\mathrm{tr}\left[\Pi_{X'}^{\rho_{X'}}\Gamma_{X'}\right]\right)^{-1}\Pi_{X'}^{\rho_{X'}}\ . \tag{C.34}$$

Then $\mathrm{tr}\,(\omega_{X'}\Gamma_{X'}) = 1$, satisfying the dual constraint (C.16a). Let's now study (C.16b):

$$\begin{aligned}\rho_{R_X}^{1/2}Z_{X'R_X}\rho_{R_X}^{1/2} &- \Gamma_{R_X}\otimes\omega_{X'}\\ &= \mu\Pi_{R_X}^{\rho_{R_X}}\,\Phi_{X':R_X}\,\Pi_{R_X}^{\rho_{R_X}} - \left(\mathrm{tr}\left[\Pi_{X'}^{\rho_{X'}}\Gamma_{X'}\right]\right)^{-1}\Gamma_{R_X}\otimes\Pi_{X'}^{\rho_{X'}}\ .\end{aligned} \tag{C.35}$$

The operator $\Pi_{R_X}^{\rho_{R_X}}\Phi_{X'R_X}\Pi_{R_X}^{\rho_{R_X}}$ is a rank-1 positive operator with support within $\Pi_{R_X}^{\rho_{R_X}}\otimes\Pi_{X'}^{\rho_{X'}}$, and its nonzero eigenvalue is given by

$$\mathrm{tr}\left(\Pi_{R_X}^{\rho_{R_X}}\Phi_{X'R}\Pi_{R_X}^{\rho_{R_X}}\right) = \mathrm{rank}\,\rho_{R_X}\ . \tag{C.36}$$

Let $r = \mathrm{rank}\,\rho_{R_X}$. We then have $\Pi_{R_X}^{\rho_{R_X}}\Phi_{X'R}\Pi_{R_X}^{\rho_{R_X}} \leqslant r\Pi_{R_X}^{\rho_{R_X}}\otimes\Pi_{X'}^{\rho_{X'}}$ and we may continue our calculation:

$$\text{(C.35)} \leqslant \left(\mu r\Pi_{R_X}^{\rho_{R_X}} - \left(\mathrm{tr}\left[\Pi_{X'}^{\rho_{X'}}\Gamma_{X'}\right]\right)^{-1}\Gamma_{R_X}\right)\otimes\Pi_{X'}^{\rho_{X'}}\ . \tag{C.37}$$

Now, let P_{R_X} be the projector onto the eigenspaces associated to the positive (or null) eigenvalues of the operator $\left(\mu r\Pi_{R_X}^{\rho_{R_X}} - \left(\mathrm{tr}\left[\Pi_{X'}^{\rho_{X'}}\Gamma_{X'}\right]\right)^{-1}\Gamma_{R_X}\right)$, and let

$$X_{R_X} = P_{R_X}\left(\mu r\Pi_{R_X}^{\rho_{R_X}} - \left(\mathrm{tr}\left[\Pi_{X'}^{\rho_{X'}}\Gamma_{X'}\right]\right)^{-1}\Gamma_{R_X}\right)P_{R_X}\ . \tag{C.38}$$

Then

$$\text{(C.37)} \leqslant X_{R_X}\otimes\mathbb{1}_{X'}\ . \tag{C.39}$$

Hence, for any $\mu > 0$, this choice of dual variables satisfies the dual constraints. The value attained by this choice of variables is given by

$$\mathrm{tr}\left[Z_{X'R}\rho_{X'R}\right] - \mathrm{tr}\,X_{R_X} = \mu\,\mathrm{tr}\left[\Pi_{R_X}^{\rho_{R_X}}\Phi_{X'R}\Pi_{R_X}^{\rho_{R_X}}\Phi_{X'R}\right] - \mathrm{tr}\,X_{R_X}\ . \tag{C.40}$$

As the object $\Pi_{R_X}^{\rho_{R_X}}\Phi_{X'R}\Pi_{R_X}^{\rho_{R_X}}$ is rank-1, we have thanks to (C.36) that $\mathrm{tr}\left[\left(\Pi_{R_X}^{\rho_{R_X}}\Phi_{X'R}\Pi_{R_X}^{\rho_{R_X}}\right)^2\right] = \left(\mathrm{tr}\,\Pi_{R_X}^{\rho_{R_X}}\Phi_{X'R}\Pi_{R_X}^{\rho_{R_X}}\right)^2 = r^2$. Then

$$\begin{aligned}\text{(C.40)} &= \mu r^2 - \mathrm{tr}\,X_{R_X}\\ &= \mu r^2 - \mu r\,\mathrm{tr}\left(P_{R_X}\Pi_{R_X}^{\rho_{R_X}}\right) + \left(\mathrm{tr}\left[\Pi_{X'}^{\rho_{X'}}\Gamma_{X'}\right]\right)^{-1}\mathrm{tr}\left(P_{R_X}\Gamma_{R_X}\right)\\ &\geqslant \mu r^2 - \mu r\,\mathrm{tr}\left(\Pi_{R_X}^{\rho_{R_X}}\right) + \left(\mathrm{tr}\left[\Pi_{X'}^{\rho_{X'}}\Gamma_{X'}\right]\right)^{-1}\mathrm{tr}\left(P_{R_X}\Gamma_{R_X}\right)\\ &\geqslant \left(\mathrm{tr}\left[\Pi_{X'}^{\rho_{X'}}\Gamma_{X'}\right]\right)^{-1}\mathrm{tr}\left(P_{R_X}\Gamma_{R_X}\right)\ ,\end{aligned} \tag{C.41}$$

recalling that $\mathrm{tr}\,\Pi_{R_X}^{\rho_{R_X}} = \mathrm{rank}\,\rho_{R_X} = r$.

Next episode: the Lemma awakens. Take $A = \mu r\Pi_{R_X}^{\rho_{R_X}}$ and $B = \left(\mathrm{tr}\left[\Pi_{X'}^{\rho_{X'}}\Gamma_{X'}\right]\right)^{-1}\Gamma_{R_X}$; Lemma 41 then asserts that there exists a constant c independent of μ such that

$$\Pi_{R_X}^{\rho_{R_X}} \leqslant P_{R_X} + \frac{c}{\mu}\mathbb{1}\ . \tag{C.42}$$

Hence,

$$\text{(C.41)} \geqslant \left(\mathrm{tr}\left[\Pi_{X'}^{\rho_{X'}}\Gamma_{X'}\right]\right)^{-1}\left(\mathrm{tr}\left[\Pi_{R_X}^{\rho_{R_X}}\Gamma_{R_X}\right] - \frac{c}{\mu}\,\mathrm{tr}\,\Gamma_{R_X}\right) = 1 - O\left(1/\mu\right)\ . \tag{C.43}$$

Taking $\mu \to \infty$ yields successive feasible dual candidates with attained objective value converging to 1, hence proving that

$$\hat{D}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) \leqslant 0\ . \qquad\blacksquare$$

An essentially trivial proposition immediately follows from the fact that Γ-sub-preserving maps are admissible operations, and hence don't cost anything in our framework:

Proposition 14. *Let σ_X be a quantum state and let $\mathcal{E}_{X\to X'}$ be a Γ-sub-preserving logical process. With the process matrix $\rho_{X'R} = \mathcal{E}_{X\to X'}(\sigma_X^{1/2}\,\Phi_{X:R_X}\,\sigma_X^{1/2})$, we have for any $\epsilon \geqslant 0$,*

$$\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) \geqslant 0\ . \tag{C.44}$$

Proof of Proposition 14. The process $\mathcal{E}_{X\to X'}$ itself is a valid optimization candidate in (C.13), and clearly $\left\|\Gamma_{X'}^{-1/2}\,\mathcal{E}_{X\to X'}(\Gamma_X)\,\Gamma_{X'}^{-1/2}\right\|_\infty \leqslant \left\|\Pi_{X'}^{\Gamma_{X'}}\right\|_\infty \leqslant 1$ because $\mathcal{E}_{X\to X'}$ is Γ-sub-preserving. ■

In general, the coherent relative entropy depends on the precise logical process used to map the input and output states. However, there are some classes of states for which the coherent relative entropy depends only on the input and output state.

The following proposition tells us that one may map the $\Gamma_X/\mathrm{tr}\,\Gamma_X$ state to the $\Gamma_{X'}/\mathrm{tr}\,\Gamma_{X'}$ state in however way one wants, i.e. regardless of the logical process, and yet in any case the coherent relative entropy is given by the ratio $\mathrm{tr}\,\Gamma_{X'}/\mathrm{tr}\,\Gamma_X$. This is a consequence of allowing any Γ-preserving maps to be performed for free, and this ratio comes about from the normalization of the respective input and output states.

Proposition 15. *Let P_X and $P'_{X'}$ be projectors with $[P_X, \Gamma_X] = 0$ and $[P'_{X'}, \Gamma_{X'}] = 0$. Let $\rho_{X'R_X}$ be a bipartite quantum state with reduced states $\rho_{R_X} = t_{X\to R_X}[(P_X\Gamma_X P_X)/\mathrm{tr}(P_X\Gamma_X)]$ and $\rho_{X'} = (P'_{X'}\Gamma_{X'}P'_{X'})/\mathrm{tr}(P'_{X'}\Gamma_{X'})$. Then, for any $\epsilon \geqslant 0$,*

$$\begin{aligned}&\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \\ &\quad = \log \mathrm{tr}(P'_{X'}\Gamma_{X'}) - \log \mathrm{tr}(P_X\Gamma_X) + \log[1/(1-\epsilon^2)] \ . \end{aligned} \tag{C.45}$$

Proof of Proposition 15. Let $|\rho\rangle_{X'R_XE}$ be a purification of $\rho_{X'R_X}$ into a (large enough) system E, and consider the semidefinite program given by Proposition 4. We give feasible primal and dual candidates which achieve the same value. First, let $T_{X'R_XE} = (1-\epsilon^2)\,\rho_{R_X}^{-1/2}\,\rho_{X'R_XE}\,\rho_{R_X}^{-1/2}$. We have $\mathrm{tr}_{X'E}(T_{X'R_XE}) = (1-\epsilon^2)\,\Pi^{\rho_{R_X}}_{R_X} \leqslant \mathbb{1}_{R_X}$ as required by (C.4a). Also, since $\rho_{R_X} = P_{R_X}\Gamma_{R_X}P_{R_X}/\mathrm{tr}(P_{R_X}\Gamma_{R_X})$ and $\rho_{X'} = P'_{X'}\Gamma_{X'}P'_{X'}/\mathrm{tr}(P'_{X'}\Gamma_{X'})$, we have $\mathrm{tr}_{R_XE}(T_{X'R_XE}\,\Gamma_{R_X}) = (1-\epsilon^2)\,\mathrm{tr}(P_{R_X}\Gamma_{R_X})\,\mathrm{tr}_{R_X}(\rho_{X'R_X}\,P_X) = (1-\epsilon^2)\,\mathrm{tr}(P_{R_X}\Gamma_{R_X})\,\rho_{X'} \leqslant \alpha\,\Gamma_{X'}$, where we have defined $\alpha = (1-\epsilon^2)\,\mathrm{tr}(P_{R_X}\Gamma_{R_X})/\mathrm{tr}(P'_{X'}\Gamma_{X'})$ and noting that $[P'_{X'}, \Gamma_{X'}] = 0$, hence satisfying (C.4b). Finally, we have $\mathrm{tr}\big[\rho_{R_X}^{1/2}\,T_{X'R_XE}\,\rho_{R_X}^{1/2}\,\rho_{X'R_XE}\big] = (1-\epsilon^2)$ which satisfies (C.4c). This choice of primal variables is feasible, and attains the value α.

Now we exhibit feasible dual candidates. Let $\mu = \mathrm{tr}(P_{R_X}\Gamma_{R_X})/\mathrm{tr}(P'_{X'}\Gamma_{X'})$, $\omega_{X'} = P'_{X'}/\mathrm{tr}(P'_{X'}\,\Gamma_{X'})$ and $X_{R_X} = 0$, and note that (C.5a) is automatically satisfied. Then, since $\rho_{X'R_XE} \leqslant \mathbb{1}_E \otimes P'_{X'} \otimes P_{R_X}$, we have

$$\begin{aligned}\mu\,\rho_{R_X}^{1/2}\,\rho_{X'R_XE}\,\rho_{R_X}^{1/2} &\leqslant \frac{\mathrm{tr}\,P_{R_X}\Gamma_{R_X}}{\mathrm{tr}\,P'_{X'}\Gamma_{X'}}\,\mathbb{1}_E \otimes P'_{X'} \otimes \rho_{R_X} \\ &\leqslant \mathbb{1}_E \otimes \omega_{X'} \otimes \Gamma_{R_X} \ ,\end{aligned} \tag{C.46}$$

keeping in mind that $[P_{R_X}, \Gamma_{R_X}] = 0$, and hence condition (C.5b) is satisfied. The value attained by this choice of variables is simply $\mu\,(1-\epsilon^2) - \mathrm{tr}\,X_{R_X} = \alpha$, hence proving that this is the optimal solution of the semidefinite program. Calculating $-\log\alpha$ completes the proof. ∎

We note that for this special type of states we have the nice expression for their relative entropy to Γ.

Proposition 16. *If $\Gamma \geqslant 0$ and P is a projector with $[P, \Gamma] = 0$, then*

$$\begin{aligned}D\left(\frac{P\Gamma P}{\mathrm{tr}\,P\Gamma}\,\middle\|\,\Gamma\right) = D_{\mathrm{min},0}\left(\frac{P\Gamma P}{\mathrm{tr}\,P\Gamma}\,\middle\|\,\Gamma\right) = D_{\mathrm{max}}\left(\frac{P\Gamma P}{\mathrm{tr}\,P\Gamma}\,\middle\|\,\Gamma\right) \\ = -\log\mathrm{tr}\,P\Gamma \ .\end{aligned} \tag{C.47}$$

Proof of Proposition 16. Write as shorthand $\rho = P\Gamma P/\mathrm{tr}\,P\Gamma$. Then

$$\begin{aligned}2^{D_{\mathrm{max}}(\rho\,\|\,\Gamma)} &= \|\Gamma^{-1/2}\,\rho\,\Gamma^{-1/2}\|_\infty \\ &= (\mathrm{tr}\,P\Gamma)^{-1}\,\|\Gamma^{-1/2}\,P\,\Gamma\,P\,\Gamma^{-1/2}\|_\infty \\ &= (\mathrm{tr}\,P\Gamma)^{-1}\,\|\Gamma^{-1/2}\,\Gamma^{1/2}\,P\,\Gamma^{1/2}\,\Gamma^{-1/2}\|_\infty \\ &= (\mathrm{tr}\,P\Gamma)^{-1} \ ,\end{aligned} \tag{C.48}$$

since $[P, \Gamma] = 0$. Also, observing that $\Pi^\rho = P$,

$$2^{-D_{\mathrm{min},0}(\rho\,\|\,\Gamma)} = \mathrm{tr}(\Pi^\rho\Gamma) = \mathrm{tr}(P\Gamma) \ . \tag{C.49}$$

The expression $D(\rho\,\|\,\Gamma)$ is thus also equal to $-\log\mathrm{tr}\,P\Gamma$ since we know that $D_{\mathrm{min},0}(\rho\,\|\,\Gamma) \leqslant D(\rho\,\|\,\Gamma) \leqslant D_{\mathrm{max}}(\rho\,\|\,\Gamma)$ [62, Lemma 10]. ∎

Notably, the states of the form $P\Gamma P/\mathrm{tr}(P\Gamma)$ for $[P, \Gamma] = 0$ are precisely those general type of states which we allowed on battery systems in item (v) of Proposition 3.

In fact, we may prove a slightly more general version of Proposition 15 for the case $\epsilon = 0$: it suffices that the reduced state on the input is of the form $\Gamma_X/\mathrm{tr}\,\Gamma_X$, and then the coherent relative entropy is oblivious to any correlation between input and output, or equivalently, to which process is exactly implemented, and depends only on the reduced states on the input and the output.

Proposition 17. *Let $\rho_{X'R_X}$ such that $\mathrm{tr}_{X'}\,\rho_{X'R_X} = \Gamma_{R_X}/\mathrm{tr}\,\Gamma_{R_X}$. Then*

$$\hat{D}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) = -\log\mathrm{tr}\,\Gamma_X - D_{\mathrm{max}}(\rho_{X'} \,\|\, \Gamma_{X'}) \ . \tag{C.50}$$

Proof of Proposition 17. Take any $T_{X'R_X}$ satisfying $\rho_{R_X}^{1/2}\,T_{X'R_X}\,\rho_{R_X}^{1/2} = \rho_{X'R_X}$ and $\mathrm{tr}_{X'}\,T_{X'R_X} \leqslant \mathbb{1}_{R_X}$. Then since $\mathrm{tr}(\Gamma_{R_X})\,\rho_{R_X} = \Gamma_{R_X}$, we have

$$\begin{aligned}\mathrm{tr}_{R_X}(T_{X'R_X}\Gamma_{R_X}) &= \mathrm{tr}(\Gamma_{R_X})\,\mathrm{tr}_{R_X}(\rho_{R_X}^{1/2}\,T_{X'R_X}\,\rho_{R_X}^{1/2}) \\ &= \mathrm{tr}(\Gamma_{R_X})\,\mathrm{tr}_{R_X}(\rho_{X'R_X}) = \mathrm{tr}\,(\Gamma_{R_X})\,\rho_{X'} \ ,\end{aligned} \tag{C.51}$$

and thus

$$\begin{aligned}-\log\big\|\Gamma_{X'}^{-1/2}\,\mathrm{tr}_{R_X}[T_{X'R_X}\Gamma_{R_X}]\,\Gamma_{X'}^{-1/2}\big\|_\infty &= -\log\big\|\Gamma_{X'}^{-1/2}\,\rho_{X'}\Gamma_{X'}^{-1/2}\big\|_\infty \\ &= -D_{\mathrm{max}}(\rho_{X'} \,\|\, \Gamma_{X'}) \ .\end{aligned} \tag{C.52}$$

This argument holds in particular for the optimal such $T_{X'R_X}$. ∎

Remarkably, if $\mathrm{tr}_{R_X}\,\rho_{X'R_X} = \Gamma_{X'}/\mathrm{tr}\,\Gamma_{X'}$, the coherent relative entropy may still depend on the exact process, and does not necessarily reduce to a difference of input and output terms as in (C.50). This can be seen by considering the unitary process $\mathcal{U}$ which swaps two levels $|0\rangle, |1\rangle$, choosing $\Gamma = g_0|0\rangle\langle 0| + g_1|1\rangle\langle 1|$ (with $g_0 + g_1 = 1$ and $g_0 > g_1$) for both input and output, and using the input state $\sigma = g_1|0\rangle\langle 0| + g_0|1\rangle\langle 1|$: in this case, σ maps to Γ, but $-\log\|\Gamma^{-1/2}\,\mathcal{U}(\Gamma)\,\Gamma^{-1/2}\|_\infty = -D_{\mathrm{max}}(\sigma\,\|\,\Gamma)$ whereas there are processes which map σ to Γ, such as $\mathcal{T}(\cdot) = \mathrm{tr}(\Pi^\sigma(\cdot))\,\Gamma$, which achieve a coherent relative entropy of $D_{\mathrm{min},0}(\sigma\,\|\,\Gamma)$.

3. Data processing inequality

The data processing inequality is an important property desirable for an information measure. Intuitively, it asserts that processing information cannot make it more "valuable."

In our case, the data processing inequality asserts that post-processing, or applying a map to both the output state and output Γ, may only increase the coherent relative entropy.

Proposition 18 (Data processing inequality)**.** *Let $\rho_{X'R_X}$ be a quantum state and let $\Gamma_X, \Gamma_{X'} \geqslant 0$. Let $\mathcal{F}_{X'\to X''}$ be a trace-preserving, completely positive map. Then, for any $\epsilon \geqslant 0$,*

$$\begin{aligned}&\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \\ &\quad \leqslant \hat{D}^{\epsilon}_{X\to X''}(\mathcal{F}_{X'\to X''}(\rho_{X'R_X}) \,\|\, \Gamma_X, \mathcal{F}_{X'\to X''}(\Gamma_{X'})) \ .\end{aligned} \tag{C.53}$$

(Proof on page 22.)

Proof of Proposition 18. Let $\mathcal{T}_{X\to X'}, y$ be optimal candidates for the optimization defining $2^{-\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}$ in (c.1). We construct an optimization candidate for the coherent relative entropy of the post-processed state. Let $\mathcal{T}'_{X\to X''} = \mathcal{F}_{X'\to X''}\circ\mathcal{T}_{X\to X'}$. Then $\mathcal{T}'^{\dagger}_{X\leftarrow X''}(\mathbb{1}_{X''}) = \mathcal{T}^{\dagger}_{X\leftarrow X'}(\mathcal{F}^{\dagger}_{X'\leftarrow X''}(\mathbb{1}_{X''})) \leqslant \mathbb{1}_{R_X}$ because $\mathcal{F}_{X'\to X''}$ is trace-preserving. Also, $\mathcal{T}'_{X\to X''}(\Gamma_X) \leqslant \alpha\,\mathcal{F}_{X'\to X''}(\Gamma_{X'})$. Finally, writing $|\sigma\rangle_{XR_X} = \rho_{R_X}^{1/2}\,|\Phi\rangle_{X:R_X}$, we have $P(\mathcal{T}'_{X\to X''}(\sigma_{XR}), \mathcal{F}_{X'\to X''}(\rho_{X'R_X})) \leqslant P(\mathcal{T}_{X\to X'}(\sigma_{XR}), \rho_{X'R_X}) \leqslant \epsilon$. ■

The case of pre-processing, i.e. when a map is applied to the input before the actual mapping is carried out, is less clear how to formulate. Indeed, the expression $\hat{D}^{\epsilon}_{R_{\tilde{X}}\to X'}(\mathcal{F}_{R_X\to R_{\tilde{X}}}(\rho_{X'R_X})\,\|\,\mathcal{F}_{X\to\tilde{X}}(\Gamma_X),\Gamma_{X'})$ would correspond to the not-so-natural setting where one implements a process matrix defined by the state resulting when two logical processes are applied on both the system X of interest and the reference system R_X on a pure state $|\sigma\rangle_{XR_X}$. However, a more general statement about composing processes can be derived in the form of a chain rule, which is the subject of the next section.

4. Chain rule

If two individual processes are concatenated, what can be said of the coherent relative entropy of the combined processes? As one would expect, it turns out that the optimal battery use of implementing directly a composition of logical maps can only be better than the sum of the battery uses of the individual realizations of each map.

Proposition 19 (Chain rule). *Consider three systems X, X', X'' with corresponding $\Gamma_X, \Gamma_{X'}, \Gamma_{X''} \geqslant 0$, and let $R_X \simeq X, R_{X'} \simeq X'$. Let σ_X be a quantum state. Let $\mathcal{E}^{(1)}_{X\to X'}$ and $\mathcal{E}^{(2)}_{X'\to X''}$ be two completely positive, trace-nonincreasing maps such that $\mathrm{tr}[\mathcal{E}^{(2)}_{X'\to X''}(\mathcal{E}^{(1)}_{X\to X'}(\sigma_X))] = 1$. Let $\epsilon,\epsilon' \geqslant 0$. Then:*

$$\begin{aligned}\hat{D}^{\epsilon}_{X\to X'}&(\mathcal{E}^{(1)}_{X\to X'}(\sigma_{XR_X})\,\|\,\Gamma_X,\Gamma_{X'})\\ &+ \hat{D}^{\epsilon'}_{X'\to X''}(\mathcal{E}^{(2)}_{X'\to X''}(\rho'_{X'R'_X})\,\|\,\Gamma_{X'},\Gamma_{X''})\\ \leqslant\ &\hat{D}^{\epsilon+\epsilon'}_{X\to X''}(\mathcal{E}^{(2)}_{X'\to X''}(\mathcal{E}^{(1)}_{X\to X'}(\sigma_{XR_X}))\,\|\,\Gamma_X,\Gamma_{X''})\,, \qquad (c.54)\end{aligned}$$

where $|\sigma\rangle_{XR_X} = \sigma_X^{1/2}\,|\Phi\rangle_{X:R_X}$ and $|\rho'\rangle_{X'R_{X'}} = (\mathcal{E}^{(1)}_{X\to X'}(\sigma_X))^{1/2}\,|\Phi\rangle_{X':R_{X'}}$.

Proof of Proposition 19. Let $\mathcal{T}^{(1)}_{X\to X'}$, y_1 be optimal choices in (c.1) for $\hat{D}^{\epsilon}_{X\to X'}(\mathcal{E}^{(1)}_{X\to X'}(\sigma_{XR_X})\,\|\,\Gamma_X,\Gamma_{X'})$, and let $\mathcal{T}^{(2)}_{X\to X'}$, y_2 be optimal choices for $\hat{D}^{\epsilon'}_{X'\to X''}(\mathcal{E}^{(2)}_{X'\to X''}(\rho'_{X'R'_X})\,\|\,\Gamma_{X'},\Gamma_{X''})$. Let $V_{X\to X'E}$ be a Stinespring dilation of $\mathcal{E}^{(1)}_{X\to X'}$, such that $\mathcal{E}^{(1)}_{X\to X'}(\cdot) = \mathrm{tr}_E[V_{X\to X'E}\,(\cdot)V^{\dagger}]$. Now, as two different purifications of $\mathcal{E}^{(1)}_{X\to X'}(\sigma_X) = \rho'_{X'}$, there must exist a partial isometry $W_{R_{X'}\to R_X E}$ such that $V_{X\to X'E}\,|\sigma\rangle_{XR_X} = W_{R_{X'}\to R_X E}|\rho'\rangle_{X'R_{X'}}$. Define $\mathcal{F}_{R_{X'}\to R_X}(\cdot) = \mathrm{tr}_E(W_{R_{X'}\to R_X E}\,(\cdot)\,W^{\dagger})$, and note that $\mathcal{E}^{(1)}_{X\to X'}(\sigma_{XR_X}) = \mathcal{F}_{R_{X'}\to R_X}(\rho'_{X'R_{X'}})$. Now, let $\mathcal{T}_{X\to X''} = \mathcal{T}^{(2)}_{X'\to X''}\circ\mathcal{T}^{(1)}_{X\to X'}$, and note that

$$\begin{aligned}P\big[&\mathcal{T}_{X\to X''}(\sigma_{XR_X}),\,\mathcal{E}^{(2)}_{X'\to X''}(\mathcal{E}^{(1)}_{X\to X'}(\sigma_{XR_X}))\big]\\ &\leqslant P\big[\mathcal{T}^{(2)}_{X'\to X''}(\mathcal{T}^{(1)}_{X\to X'}(\sigma_{XR_X})),\,\mathcal{T}^{(2)}_{X'\to X''}(\mathcal{E}^{(1)}_{X\to X'}(\sigma_{XR_X}))\big]\\ &\quad + P\big[\mathcal{T}^{(2)}_{X'\to X''}(\mathcal{E}^{(1)}_{X\to X'}(\sigma_{XR_X})),\,\mathcal{E}^{(2)}_{X'\to X''}(\mathcal{E}^{(1)}_{X\to X'}(\sigma_{XR_X}))\big]\\ &\leqslant P\big[\mathcal{T}^{(1)}_{X\to X'}(\sigma_{XR_X}),\,\mathcal{E}^{(1)}_{X\to X'}(\sigma_{XR_X})\big]\\ &\quad + P\big[\mathcal{T}^{(2)}_{X'\to X''}(\rho'_{X'R_{X'}}),\,\mathcal{E}^{(2)}_{X'\to X''}(\rho'_{X'R_{X'}})\big]\\ &\leqslant \epsilon+\epsilon'\,. \qquad (c.55)\end{aligned}$$

where in second inequality we have used twice the fact that the purified distance cannot decrease under application of a completely positive, trace-nonincreasing map, and that $\mathcal{E}^{(1)}_{X\to X'}(\sigma_{XR_X}) = \mathcal{F}_{R_{X'}\to R_X}(\rho'_{X'R_{X'}})$. Observe finally that

$$\begin{aligned}\mathcal{T}_{X\to X''}(\Gamma_X) = \mathcal{T}^{(2)}_{X'\to X''}(\mathcal{T}^{(1)}_{X\to X'}(\Gamma_X)) &\leqslant 2^{-y_1}\,\mathcal{T}^{(2)}_{X'\to X''}(\Gamma_{X'})\\ &\leqslant 2^{-y_1-y_2}\,\Gamma_{X''}\,, \qquad (c.56)\end{aligned}$$

proving that $\mathcal{T}_{X\to X''}$, $y = y_1+y_2$ are valid optimization candidates in (c.1) for $\hat{D}^{\epsilon+\epsilon'}_{X\to X''}(\mathcal{E}^{(2)}_{X'\to X''}(\mathcal{E}^{(1)}_{X\to X'}(\sigma_{XR_X}))\,\|\,\Gamma_X,\Gamma_{X''})$, proving the claim. ■

Corollary 20 (Chain rule in terms of states). *Consider systems A, B, C and $R_A \simeq A$, $R_B \simeq B$. Let $\Gamma_C \geqslant 0$, $\Gamma_{AB} \geqslant 0$ and write $\Gamma_A = \mathrm{tr}_B[\Gamma_{AB}]$. Let $\tau_{CR_AR_B}$ be any tripartite state. Then, for $\epsilon,\epsilon' \geqslant 0$,*

$$\begin{aligned}\hat{D}^{\epsilon}_{A\to AB}(\rho_{ABR_A}\,\|\,\Gamma_A,\Gamma_{AB}) + \hat{D}^{\epsilon'}_{AB\to C}(\tau_{CR_AR_B}\,\|\,\Gamma_{AB},\Gamma_C)&\\ \leqslant \hat{D}^{\epsilon+\epsilon'}_{A\to C}(\tau_{CR_A}\,\|\,\Gamma_A,\Gamma_C)\,,& \qquad (c.57)\end{aligned}$$

where $\rho_{ABR_A} = \mathrm{tr}_{R_B}[\tau^{1/2}_{R_AR_B}\,\Phi_{AB:R_AR_B}\,\tau^{1/2}_{R_AR_B}]$.

Proof of Corollary 20. Define systems $X = A$, $X' = AB$ and $X'' = C$. Let

$$\mathcal{E}^{(1)}_{X\to X'}(\cdot) = \mathrm{tr}_{R_A}[\rho_{R_A}^{-1/2}\,\rho_{ABR_A}\,\rho_{R_A}^{-1/2}\,t_{A\to R_A}(\cdot)]\,; \qquad (c.58a)$$

$$\mathcal{E}^{(2)}_{X'\to X''}(\cdot) = \mathrm{tr}_{R_AR_B}[\tau^{-1/2}_{R_AR_B}\,\tau_{CR_AR_B}\,\tau^{-1/2}_{R_AR_B}\,t_{AB\to R_AR_B}(\cdot)]\,. \qquad (c.58b)$$

These mappings are trace nonincreasing. Let $\sigma_X = t_{R_X\to X}(\tau_{R_X}) = t_{R_X\to X}(\rho_{R_X})$. We see that $\mathcal{E}^{(2)}_{X'\to X''}(\mathcal{E}^{(1)}_{X\to X'}(\sigma_X)) = \mathcal{E}^{(2)}_{X'\to X''}(\rho_{AB}) = \mathcal{E}^{(2)}_{X'\to X''}(t_{R_AR_B\to AB}(\tau_{R_AR_B})) = \tau_C$ which has unit trace as required. Furthermore, let $|\sigma\rangle_{XR_X} = \sigma_X^{1/2}\,|\Phi\rangle_{X:R_X} = \sigma_A^{1/2}\,|\Phi\rangle_{A:R_A}$ and $|\rho'\rangle_{X'R_{X'}} = (\rho_{AB}^{1/2})\,|\Phi\rangle_{AB:R_AR_B} = (\tau^{1/2}_{R_AR_B})\,|\Phi\rangle_{ABR_AR_B}$. Now calculate

$$\begin{aligned}\mathcal{E}^{(1)}_{X\to X'}(\sigma_{XR_X}) = \Pi^{\rho_{R_A}}_{R_A}\,\mathrm{tr}_{\bar{R}_A}[\rho_{AB\bar{R}_A}\,t_{A\to\bar{R}_A}(\Phi_{A:R_A})]\,\Pi^{\rho_{R_A}}_{R_A}&\\ = \rho_{ABR_A}\,,& \qquad (c.59)\end{aligned}$$

as well as

$$\begin{aligned}&\mathcal{E}^{(2)}_{X'\to X''}(\rho'_{X'R_{X'}}) =\\ &\quad\Pi^{\tau_{R_AR_B}}_{R_AR_B}\,\mathrm{tr}_{\bar{R}_A\bar{R}_B}[\tau_{C\bar{R}_A\bar{R}_B}\,t_{AB\to\bar{R}_A\bar{R}_B}(\Phi_{AB:R_AR_B})]\,\Pi^{\tau_{R_AR_B}}_{R_AR_B}\\ &\qquad = \tau_{CR_AR_B}\,, \qquad (c.60)\end{aligned}$$

and, since $\mathcal{E}^{(1)}_{X\to X'}(\sigma_{XR_X}) = \rho_{ABR_A} = \mathrm{tr}_{R_B}[\rho'_{ABR_AR_B}]$,

$$\begin{aligned}&\mathcal{E}^{(2)}_{X'\to X''}(\mathcal{E}^{(1)}_{X\to X'}(\sigma_{XR_X}))\\ &\quad = \mathrm{tr}_{R_B}\big[\mathcal{E}^{(2)}_{X'\to X''}(\rho'_{ABR_AR_B})\big]\\ &\quad = \tau_{CR_A}\,. \qquad (c.61)\end{aligned}$$

All conditions for Proposition 19 are fulfilled, and the claim follows. ■

5. Alternative smoothing of the coherent relative entropy

There is another possible way to define the smooth coherent relative entropy (i.e., for $\epsilon > 0$), based on optimizing its non-smooth version (for $\epsilon = 0$) over all states which are ϵ-close to the requested state. This smoothing method is the method used traditionally in the smooth entropy framework [61, 63, 69]. The disadvantage of this alternative definition is that it can no longer be formulated as a semidefinite program. However, in the regime of small ϵ, it turns out that both definitions are equivalent up to factors which depend only on ϵ, and which do not scale with the dimension of the system (Proposition 25 below). In particular, both quantities behave in the same way in the i.i.d. limit.

Alternative smoothing. *For a normalized state $\rho_{X'R_X}$, positive semidefinite $\Gamma_X, \Gamma_{X'}$, and for $\epsilon \geqslant 0$, we define the quantity*

$$\bar{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) = \max_{\hat\rho_{X'R_X} \approx_\epsilon \rho_{X'R_X}} \hat{D}_{X\to X'}(\hat\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \,, \tag{C.62}$$

where the maximization in (C.62) is taken over (normalized) quantum states which are in the support of $\Gamma_X \otimes \Gamma_{X'}$ and which are close to $\rho_{X'R_X}$ in the purified distance, $P(\hat\rho_{X'R_X}, \rho_{X'R_X}) \leqslant \epsilon$.

Some properties of $\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'})$ carry over immediately to $\bar{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'})$, which we summarize here without explicit proof. These propositions are straightforwardly proven by applying the relevant property to the inner coherent relative entropy in (C.62).

Proposition 21 (cf. Proposition 5)**.** *For any $0 \leqslant \epsilon < 1$,*

$$\bar{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \geqslant -\log \operatorname{tr}\Gamma_X - \log \|\Gamma_{X'}^{-1}\|_\infty \,; \tag{C.63a}$$
$$\bar{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \leqslant \log \|\Gamma_X^{-1}\|_\infty + \log \operatorname{tr}\Gamma_{X'} \,. \tag{C.63b}$$

Proposition 22 (cf. Proposition 7)**.** *For any $a, b \geqslant 0$,*

$$\bar{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, a\Gamma_X, b\Gamma_{X'}) = \bar{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) + \log\frac{b}{a} \,. \tag{C.64}$$

Proposition 23 (cf. Proposition 6)**.** *Let $\tilde{X}$, $\tilde{X}'$ be new systems. Suppose there exist partial isometries $V_{X\to\tilde{X}}$ and $V'_{X'\to\tilde{X}'}$ such that both $t_{R_X\to X}(\rho_{R_X})$ and Γ_X are in the support of $V_{X\to\tilde{X}}$, and both $\rho_{X'}$ and $\Gamma_{X'}$ are in the support of $V'_{X'\to\tilde{X}'}$. Then*

$$\bar{D}^{\epsilon}_{\tilde{X}\to\tilde{X}'}((V'\otimes V)\,\rho_{X'R_X}\,(V'\otimes V)^\dagger \,\|\, V\Gamma_X V^\dagger, V'\Gamma_{X'}V'^\dagger) = \bar{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \,. \tag{C.65}$$

We now give a loose equivalent of Proposition 15 for the alternative smoothing of the coherent relative entropy. The error term is relatively loose (it scales proportionally to n and to ϵ), and it does not disappear in the i.i.d. limit unless the limit $\epsilon \longrightarrow 0$ is taken explicitly. For this reason, for small ϵ, it might be advantageous to use Proposition 15 in conjunction with Proposition 25.

Proposition 24. *Let $P_X, P'_{X'}$ be projectors such that $[\Gamma_X, P_X] = 0$ and $[\Gamma_{X'}, P'_{X'}] = 0$. Let $\rho_{X'R_X}$ be such that $\rho_{R_X} = t_{X\to R_X}(P_X\Gamma_X P_X/\operatorname{tr} P_X\Gamma_X)$ and $\rho_{X'} = P'_{X'}\Gamma_{X'}P'_{X'}/\operatorname{tr} P'_{X'}\Gamma_{X'}$. Let $\epsilon \geqslant 0$. Then*

$$\bar{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) = \log\frac{\operatorname{tr} P'_{X'}\Gamma_{X'}}{\operatorname{tr} P_X\Gamma_X} + f(\epsilon, \Gamma_X, \Gamma_{X'}) \,, \tag{C.66}$$

where the error term $f(\epsilon, \Gamma_X, \Gamma_{X'})$ is bounded as

$$0 \leqslant f(\epsilon, \Gamma_X, \Gamma_{X'}) \leqslant f_0(\epsilon, \Gamma_X) + f_0(\epsilon, \Gamma_{X'}) \,, \tag{C.67}$$

where $f_0(\epsilon, \Gamma) = \epsilon\log(\operatorname{rank}\Gamma - 1) + \epsilon\|\log\Gamma\|_\infty + h(\epsilon)$ with the binary entropy $h(\epsilon) = -\epsilon\log\epsilon - (1-\epsilon)\log(1-\epsilon)$.

Proof of Proposition 24. The lower bound is given simply as

$$\bar{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \geqslant \bar{D}^{\epsilon=0}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) = \log\frac{\operatorname{tr} P'_{X'}\Gamma_{X'}}{\operatorname{tr} P_X\Gamma_X} \,, \tag{C.68}$$

where the latter expression is provided by Proposition 15, recalling that for $\epsilon = 0$ both versions of the smooth coherent relative entropy coincide exactly. For the upper bound, let $\hat\rho_{X'R_X}$ be the optimal state such that $P(\hat\rho_{X'R_X}, \rho_{X'R_X}) \leqslant \epsilon$ and

$$\bar{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) = \bar{D}_{X\to X'}(\hat\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \,, \tag{C.69}$$

and invoke Proposition 30 to get

$$\text{(C.69)} \leqslant D(\hat\rho_X \,\|\, \Gamma_X) - D(\hat\rho_{X'} \,\|\, \Gamma_{X'}) \,. \tag{C.70}$$

We have $D(\hat\rho_{R_X}, \rho_{R_X}) \leqslant P(\hat\rho_{R_X}, \rho_{R_X}) \leqslant \epsilon$ and analogously $D(\hat\rho_{X'}, \rho_{X'}) \leqslant \epsilon$. By continuity of the relative entropy given in Lemma 46, we get

$$\left|D(\hat\rho_{R_X} \,\|\, \Gamma_{R_X}) - D(\rho_{R_X} \,\|\, \Gamma_{R_X})\right| \leqslant f_0(\epsilon, \Gamma_{R_X}) \,;$$
$$\left|D(\hat\rho_{X'} \,\|\, \Gamma_{X'}) - D(\rho_{X'} \,\|\, \Gamma_{X'})\right| \leqslant f_0(\epsilon, \Gamma_{X'}) \,, \tag{C.71a}$$

where $f_0(\epsilon, \Gamma)$ is as given in the claim. On the other hand,

$$D(\rho_{R_X} \,\|\, \Gamma_{R_X}) - D(\rho_{X'} \,\|\, \Gamma_{X'}) = \log\operatorname{tr} P'_{X'}\Gamma_{X'} - \log\operatorname{tr} P_{R_X}\Gamma_{R_X} \,, \tag{C.72}$$

because $\rho_{R_X} = P_{R_X}\Gamma_{R_X}P_{R_X}/\operatorname{tr} P_{R_X}\Gamma_{R_X}$ and $\rho_{X'} = P'_{X'}\Gamma_{X'}P'_{X'}/\operatorname{tr} P'_{X'}\Gamma_{X'}$, as given by (C.47). This means that

$$\text{(C.70)} \leqslant \log\operatorname{tr}\frac{P'_{X'}\Gamma_{X'}}{P_{R_X}\Gamma_{R_X}} + f_0(\epsilon, \Gamma_{R_X}) + f_0(\epsilon, \Gamma_{X'}) \,. \qquad \blacksquare$$

Crucially, this alternative smoothing method does not alter the quantity much in the regime of small ϵ. In fact, both versions of the smooth coherent relative entropy are related by a simple adjustment of the ϵ parameter, and up to an error term which depends only on ϵ and doesn't scale with the system size.

Proposition 25. *Let $\rho_{X'R_X}$ be any quantum state. Then for any $\epsilon \geqslant 0$ with $3\sqrt{\epsilon} < 1$,*

$$\bar{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \leqslant \hat{D}^{3\sqrt{\epsilon}}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \,. \tag{C.73}$$

Conversely, for any $\epsilon > 0$ with $9\epsilon^{1/4} < 1$,

$$\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \leqslant \bar{D}^{9\epsilon^{1/4}}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) + \log(1/\epsilon) \,. \tag{C.74}$$

(Proof on page 24.)

We need to prove the following lemma first.

Lemma 26. *Let $\Gamma_X, \Gamma_{X'} \geqslant 0$. Let $\mathcal{T}_{X\to X'}$ be a completely positive, trace-nonincreasing map. Let $Q_X = \mathcal{T}^\dagger(\mathbb{1}_{X'})$. Assume that the support of Q_X lies within the support of Γ_X, and that $\mathcal{T}_{X\to X'}(\Gamma_X)$ lies within the support of $\Gamma_{X'}$. Then*

$$\min\left\{\alpha :\ \mathcal{T}_{X\to X'}(\Gamma_X) \leqslant \alpha\,\Gamma_{X'}\right\} \geqslant \frac{\operatorname{tr}(Q_X\Gamma_X)}{\operatorname{tr}\Gamma_{X'}} . \tag{C.75}$$

Proof of Lemma 26. The optimal α is given by

$$\begin{aligned}\alpha &= \left\|\Gamma_{X'}^{-1/2}\,\mathcal{T}_{X\to X'}(\Gamma_X)\,\Gamma_{X'}^{-1/2}\right\|_\infty\\ &\geqslant \operatorname{tr}\left[\left(\frac{\Gamma_{X'}}{\operatorname{tr}\Gamma_{X'}}\right)\Gamma_{X'}^{-1/2}\,\mathcal{T}_{X\to X'}(\Gamma_X)\,\Gamma_{X'}^{-1/2}\right]\\ &= (\operatorname{tr}\Gamma_{X'})^{-1}\ \operatorname{tr}[\mathcal{T}_{X\to X'}(\Gamma_X)] = (\operatorname{tr}\Gamma_{X'})^{-1}\ \operatorname{tr}[Q_X\Gamma_X]\,,\end{aligned} \tag{C.76}$$

where we have used that $\|\cdot\|_\infty = \max_\gamma \operatorname{tr}[\gamma\,(\cdot)]$ with γ ranging over all density operators. ∎

Proof of Proposition 25. First we prove (C.73). Let $\bar\rho_{X'R}$ be the state which achieves the optimum in $\bar D^{\epsilon}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X, \Gamma_{X'})$, and let $T_{X'R_X}$, α be optimal primal variables for $2^{-\bar D_{X\to X'}(\bar\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}$ for the semidefinite program in Proposition 12, and denote by $\mathcal{T}_{X\to X'}$ the completely positive, trace-nonincreasing map corresponding to $T_{X'R_X}$. Write $|\sigma\rangle_{XR_X} = \rho_{R_X}^{1/2}\,|\Phi\rangle_{X:R_X}$ and $|\bar\sigma\rangle_{XR_X} = \bar\rho_{R_X}^{1/2}\,|\Phi\rangle_{X:R_X}$. Since $P(\sigma_{R_X}, \bar\sigma_{R_X}) \leqslant \epsilon$, we see using Lemma 45 that $P(\sigma_{XR_X}, \bar\sigma_{XR_X}) \leqslant 2\sqrt{\epsilon}$. The purified distance may not increase under the action of the trace nonincreasing map $\mathcal{T}_{X\to X'}$, and hence

$$\begin{aligned}&P\big(\mathcal{T}_{X\to X'}(\sigma_{XR_X}), \rho_{X'R_X}\big)\\ &\quad\leqslant P\big(\mathcal{T}_{X\to X'}(\sigma_{XR_X}), \bar\rho_{X'R_X}\big) + P(\bar\rho_{X'R_X}, \rho_{X'R_X})\\ &\quad\leqslant P\big(\mathcal{T}_{X\to X'}(\sigma_{XR_X}), \mathcal{T}_{X\to X'}(\bar\sigma_{XR_X})\big) + \epsilon\\ &\quad\leqslant 2\sqrt{\epsilon} + \epsilon \leqslant 3\sqrt{\epsilon}\,.\end{aligned} \tag{C.77}$$

Hence, $\mathcal{T}_{X\to X'}$ is an optimization candidate for $2^{-\hat D^{3\sqrt{\epsilon}}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}$ with the same achieved value, proving (C.73).

Now we prove (C.74). In the remainder of this proof, we use the shorthand system name $R \equiv R_X$. Let $\hat T_{X'RE}$, $\hat\alpha$ be the optimal primal variables for $2^{-\hat D^{\epsilon}_{X\to X'}(\rho_{X'R}\,\|\,\Gamma_X,\Gamma_{X'})}$. We will construct an explicit $\tilde\rho_{X'R}$ close to $\rho_{X'R}$, as well as feasible candidates $\tilde T_{X'R}$ and $\tilde\alpha$ in the optimization for $\bar D_{X\to X'}(\tilde\rho_{X'R}\,\|\,\Gamma_X, \Gamma_{X'})$ as given by Proposition 12. We denote by $\hat{\mathcal{T}}_{X\to X'}$ the completely positive, trace nonincreasing map corresponding to $\hat T_{X'RE}$. Let $\sigma_{XR} = \rho_R^{1/2}\,\Phi_{X:R}\,\rho_R^{1/2}$ and define

$$\hat\rho_{X'R} = \hat{\mathcal{T}}_{X\to X'}(\sigma_{XR})\,. \tag{C.78}$$

By assumption, $P(\hat\rho_{X'R}, \rho_{X'R}) \leqslant \epsilon$ and hence $D(\hat\rho_{X'R}, \rho_{X'R}) \leqslant \epsilon$. Using the fact that $\hat\rho_{X'R} = \rho_{X'R} + \Delta^+_{X'R} - \Delta^-_{X'R}$ for some $\Delta^\pm_{X'R} \geqslant 0$ with $\operatorname{tr}\Delta^+_{X'R} = \operatorname{tr}\Delta^-_{X'R} = D(\hat\rho_{X'R}, \rho_{X'R}) \leqslant \epsilon$, we see that $\operatorname{tr}\hat\rho_{X'R} \geqslant \operatorname{tr}\rho_{X'R} - \epsilon = 1-\epsilon$.

Define $Q = \hat{\mathcal{T}}^\dagger(\mathbb{1}_{X'})$ and note that $0 \leqslant Q \leqslant \mathbb{1}$. For any $0 < \eta < 1$, let P^η be the projector onto the eigenspaces of Q for which the corresponding eigenvalues are greater or equal to η; clearly P^η and Q commute. Define $R^\eta = P^\eta - P^\eta Q P^\eta$, noting that P^η, Q, R^η all commute. By definition, $\eta P^\eta \leqslant P^\eta Q P^\eta$, and hence $R^\eta \leqslant (\eta^{-1}-1)\,P^\eta Q P^\eta \leqslant (\eta^{-1}-1)\,Q$. We may now define

$$\tilde{\mathcal{T}}_{X\to X'}(\cdot) = \hat{\mathcal{T}}_{X\to X'}(\cdot) + \operatorname{tr}(R^\eta\,(\cdot))\,\frac{\Gamma_{X'}}{\operatorname{tr}\Gamma_{X'}}\,. \tag{C.79}$$

The mapping $\tilde{\mathcal{T}}_{X\to X'}$ is trace non-increasing,

$$\tilde{\mathcal{T}}^\dagger_{X\leftarrow X'}(\mathbb{1}_{X'}) = Q + R^\eta = P^\eta + P^{\eta,\perp}\,Q\,P^{\eta,\perp} \leqslant \mathbb{1}\,, \tag{C.80}$$

where $P^{\eta,\perp} = \mathbb{1} - P^\eta$, keeping in mind that $Q = P^\eta Q P^\eta + P^{\eta,\perp} Q P^{\eta,\perp}$ and that $R^\eta + P^\eta Q P^\eta = P^\eta$. Furthermore $\tilde{\mathcal{T}}_{X\to X'}$ is trace-preserving on the subspace spanned by P^η, i.e. $P^\eta\,\tilde{\mathcal{T}}^\dagger_{X\leftarrow X'}(\mathbb{1}_{X'})\,P^\eta = P^\eta$. This means that for any state τ lying in the support of P^η, it holds that $\operatorname{tr}[\tilde{\mathcal{T}}_{X\to X'}(\tau)] = 1$. The map $\tilde{\mathcal{T}}_{X\to X'}$ moreover satisfies

$$\begin{aligned}\tilde{\mathcal{T}}_{X\to X'}(\Gamma_X) &\leqslant \hat\alpha\,\Gamma_{X'} + \frac{\operatorname{tr}R^\eta\Gamma_X}{\operatorname{tr}\Gamma_{X'}}\,\Gamma_{X'}\\ &\leqslant \left(\hat\alpha + (\eta^{-1}-1)\,\frac{\operatorname{tr}Q\Gamma_X}{\operatorname{tr}\Gamma_{X'}}\right)\Gamma_{X'} \leqslant \eta^{-1}\hat\alpha\,\Gamma_{X'}\,,\end{aligned} \tag{C.81}$$

where in the last inequality we have used Lemma 26 to see that $\hat\alpha \geqslant \operatorname{tr}(Q\Gamma_X)/\operatorname{tr}\Gamma_{X'}$. We are led to define (surprise!) $\tilde\alpha = \eta^{-1}\hat\alpha$.

It remains to find a state $\tilde\rho_{X'R}$ which is close to $\rho_{X'R}$ such that $\tilde\rho_R^{1/2}\,\tilde{\mathcal{T}}_{X\to X'}(\Phi_{X:R})\,\tilde\rho_R^{1/2} = \tilde\rho_{X'R}$. First define

$$\tilde\sigma_X = \frac{P^\eta\,\sigma_X\,P^\eta}{\operatorname{tr}(P^\eta\,\sigma_X)}\,. \tag{C.82}$$

Observe that $\operatorname{tr}(P^\eta\sigma_X) \geqslant \operatorname{tr}(P^\eta Q P^\eta\sigma_X) = \operatorname{tr}(Q\,\sigma_X) - \operatorname{tr}\big(P^{\eta,\perp}\,Q\,P^{\eta,\perp}\sigma_X\big) \geqslant 1-\epsilon-\eta$, where $P^{\eta,\perp} = \mathbb{1} - P^\eta$, using the fact that all eigenvalues of Q within $P^{\eta,\perp}$ are less than η and that $\operatorname{tr}(Q\,\sigma_X) = \operatorname{tr}\big(\hat{\mathcal{T}}(\sigma_X)\big) = \operatorname{tr}\hat\rho_{X'} \geqslant 1-\epsilon$. Then, using Lemma 43,

$$P(\tilde\sigma_X, \sigma_X) \leqslant \frac{\sqrt{2(\epsilon+\eta)}}{\sqrt{1-\epsilon-\eta}} =: \bar\epsilon\,. \tag{C.83}$$

Write $\tilde\sigma_{XR} = \tilde\sigma_X^{1/2}\,\Phi_{X:R}\,\tilde\sigma_X^{1/2}$. Using Lemma 45 we see that $P(\tilde\sigma_{XR}, \sigma_{XR}) \leqslant 2\sqrt{D(\tilde\sigma_R, \rho_R)} \leqslant 2\sqrt{P(\tilde\sigma_R, \rho_R)} \leqslant 2\sqrt{\bar\epsilon}$. At this point, define

$$\tilde\rho_{X'R} = \tilde{\mathcal{T}}_{X\to X'}(\tilde\sigma_{XR})\,; \tag{C.84a}$$

$$\bar\rho_{X'R} = \tilde{\mathcal{T}}_{X\to X'}(\sigma_{XR})\,. \tag{C.84b}$$

Because $\tilde\sigma_X$ lies within the support of P^η, we have $\operatorname{tr}_{X'}\tilde\rho_{X'R} = \operatorname{tr}_X\big(\tilde{\mathcal{T}}^\dagger(\mathbb{1}_{X'})\,\tilde\sigma_{XR}\big) = \operatorname{tr}_X\big(\tilde{\mathcal{T}}^\dagger(\mathbb{1}_{X'})\,P^\eta\,\tilde\sigma_{XR}\,P^\eta\big) = \tilde\sigma_R$, and hence we have $\tilde\rho_R^{1/2}\,\tilde{\mathcal{T}}_{X\to X'}(\Phi_{X:R})\,\tilde\rho_R^{1/2} = \tilde\rho_{X'R}$ as required. Furthermore, the purified distance cannot increase under the action of $\tilde{\mathcal{T}}_{X\to X'}$, so we have $P(\tilde\rho_{X'R}, \bar\rho_{X'R}) \leqslant 2\sqrt{\bar\epsilon}$. Also, $\bar\rho_{X'R} = \hat{\mathcal{T}}_{X\to X'}(\sigma_{XR}) + D_{X'R} = \hat\rho_{X'R} + D_{X'R}$ with $D_{X'R} = \operatorname{tr}(R^\eta\sigma_{XR})\,(\operatorname{tr}\Gamma_{X'})^{-1}\,\Gamma_{X'}$, noting that $\operatorname{tr}D_{X'R} \leqslant \operatorname{tr}(\bar\rho_{X'R}) - \operatorname{tr}(\hat\rho_{X'R}) \leqslant 1-(1-\epsilon) \leqslant \epsilon$; hence $D(\bar\rho_{X'R}, \hat\rho_{X'R}) \leqslant \epsilon$ and thus $P(\bar\rho_{X'R}, \hat\rho_{X'R}) \leqslant \sqrt{2\epsilon}$. We deduce that $P(\tilde\rho_{X'R}, \rho_{X'R}) \leqslant P(\tilde\rho_{X'R}, \bar\rho_{X'R}) + P(\bar\rho_{X'R}, \hat\rho_{X'R}) + P(\hat\rho_{X'R}, \rho_{X'R}) \leqslant 2\sqrt{\bar\epsilon} + \sqrt{2\epsilon} + \epsilon$.

Let's summarize: We now have a state $\tilde\rho_{X'R}$ satisfying $P(\tilde\rho_{X'R}, \rho_{X'R}) \leqslant 2\sqrt{\bar\epsilon} + \sqrt{2\epsilon} + \epsilon$, as well as a trace-nonincreasing map $\tilde{\mathcal{T}}_{X\to X'}$ satisfying $\tilde\rho_R^{1/2}\,\tilde{\mathcal{T}}_{X\to X'}(\Phi_{X:R})\,\tilde\rho_R^{1/2} = \tilde\rho_{X'R}$ and $\tilde{\mathcal{T}}_{X\to X'}(\Gamma_X) \leqslant \alpha\,\eta^{-1}\,\Gamma_{X'}$. The claim follows by choosing $\eta = \epsilon$ and calculating the bounds $\bar\epsilon \leqslant \sqrt{8\epsilon}$ (using the assumption $\epsilon < 1/4$) as well as $2\sqrt{\bar\epsilon} + \sqrt{2\epsilon} + \epsilon \leqslant \big(4\sqrt{2} + \sqrt{2} + 1\big)\,\epsilon^{1/4} \leqslant 9\,\epsilon^{1/4}$. ∎

6. Recovering known entropy measures

An interesting aspect of the coherent relative entropy is that it reduces to various previously-known entropy measures, including the min- and max-relative entropies [62], as well as the conditional min- and max-entropy [61, 69]. These measures are already known to be relevant in counting the work cost of specific processes in quantum thermodynamics [25, 30, 31, 33, 60].

First we present some definitions. Given a (normalized) quantum state ρ_{AB}, we define the *(conditional) von Neumann entropy*, the *(conditional alternative) max-entropy*, and the *(conditional al-*

ternative) min-entropy respectively as,[4]

$$
\begin{aligned}
H(A\,|\,B)_\rho &= -\operatorname{tr}(\rho_{AB}\log\rho_{AB}) + \operatorname{tr}(\rho_B\log\rho_B)\,;\\
H_{\max,0}(A\,|\,B)_\rho &= \log\big\|\operatorname{tr}_A \Pi_{AB}^{\rho_{AB}}\big\|_\infty\,;\ \text{and}\\
H_{\min,0}(A\,|\,B)_\rho &= -\log\big\|\rho_B^{-1/2}\rho_{AB}\rho_B^{-1/2}\big\|_\infty\,.
\end{aligned}
$$

For any $\epsilon > 0$, we define the *smooth (conditional alternative) max-entropy* and *smooth (conditional alternative) min-entropy* respectively as

$$
\begin{aligned}
H^\epsilon_{\max,0}(A\,|\,B)_\rho &= \min_{\hat\rho_{AB}\approx_\epsilon\rho_{AB}} H_{\max,0}(A\,|\,B)_{\hat\rho}\,;\\
H^\epsilon_{\min,0}(A\,|\,B)_\rho &= \max_{\hat\rho_{AB}\approx_\epsilon\rho_{AB}} H_{\min,0}(A\,|\,B)_{\hat\rho}\,,
\end{aligned}
$$

where the optimizations range over (normalized[5]) states $\hat\rho_{AB}$ and where $\hat\rho_{AB}\approx_\epsilon\rho_{AB}$ denotes proximity in the purified distance, i.e., $P(\hat\rho_{AB},\rho_{AB})\leqslant\epsilon$.

For a (normalized) quantum state ρ_X, and any $\Gamma_X\geqslant 0$, we define the *quantum relative entropy*, the *relative min-entropy*, and the *relative max-entropy* respectively as,

$$
\begin{aligned}
D(\rho_X\,\|\,\Gamma_X) &= \operatorname{tr}\big[\rho_X\big(\log_2\rho_X - \log_2\Gamma_X\big)\big]\,;\\
D_{\min,0}(\rho_X\,\|\,\Gamma_X) &= -\log\operatorname{tr}\big[\Pi_X^{\rho_X}\Gamma_X\big]\,;\\
D_{\max}(\rho_X\,\|\,\Gamma_X) &= \log\|\Gamma_X^{-1/2}\rho_X\Gamma_X^{-1/2}\|_\infty\,,
\end{aligned}
$$

recalling that $\Pi_X^{\rho_X}$ denotes the projector onto the support of ρ_X. We define the smoothed versions of the relative min- and max-entropies as

$$
\begin{aligned}
D^\epsilon_{\min,0}(\rho\,\|\,\Gamma) &= \max_{\hat\rho\approx_\epsilon\rho} D_{\min,0}(\hat\rho\,\|\,\Gamma)\,;\\
D^\epsilon_{\max}(\rho\,\|\,\Gamma) &= \min_{\hat\rho\approx_\epsilon\rho} D_{\max}(\hat\rho\,\|\,\Gamma)\,.
\end{aligned}
$$

where the optimizations range over normalized[6] states $\hat\rho_{AB}$ such that $P(\hat\rho_{AB},\rho_{AB})\leqslant\epsilon$.

We furthermore define the *hypothesis testing relative entropy* [73–77, 106] for $0<\eta\leqslant 1$ as

$$
D^\eta_{\mathrm{H}}(\rho\,\|\,\Gamma) = -\frac{1}{\eta}\log\min_{\substack{0\leqslant Q\leqslant\mathbb{1}\\ \operatorname{tr}[Q\rho]\geqslant\eta}}\operatorname{tr}[Q\Gamma]\,.
$$

We now show that we can recover the max-entropy in the case where for both input and output systems we have $\Gamma=\mathbb{1}$.

[4] There exist several different variants of the min- and max-entropy [61, 69]; however, all the max-entropies as well as all the min-entropies are equivalent up to terms of order $\log\epsilon$ after smoothing with a parameter ϵ.

[5] One easily notices that the normalization of the state doesn't affect these quantities, so smoothing may be restricted to normalized states (in contrast to, e.g., Refs. [63, 69]).

[6] These smooth quantities were introduced in Ref. [62] using the trace distance and optimizing over subnormalized states. The two distances are tightly related and a simple adjustment of the ϵ parameter is required. Furthermore we restrict to normalized states for our convenience; the $D^\epsilon_{\min,0}$ is not affected and the $D^\epsilon_{\max}$ is at most shifted by a factor depending on $\log(1-\epsilon)$ only.

Proposition 27 (Recovering the max-entropy)**.** *Let $|\rho\rangle_{X'R_XE}$ be any pure state on systems R_X, X', and E with $|E|\geqslant|X'R_X|$. Then*

$$
\begin{aligned}
&\bar D^\epsilon_{X\to X'}(\rho_{X'R_X}\,\|\,\mathbb{1}_X,\mathbb{1}_{X'})\\
&\qquad = -H^\epsilon_{\max,0}(E\,|\,X')_\rho = H^\epsilon_{\min,0}(E\,|\,R_X)_\rho\,. \qquad \text{(C.85)}
\end{aligned}
$$

Proof of Proposition 27. Let $|\tilde\rho\rangle_{X'R_XE}$ be any pure quantum state. Considering the semidefinite problem for $2^{-\hat D_{X\to X'}(\tilde\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}$, let $T_{X'RE} = \tilde\rho_{R_X}^{-1/2}\tilde\rho_{X'R_XE}\tilde\rho_{R_X}^{-1/2}$. Conditions (C.4a) and (C.4c) are automatically satisfied. Choosing $\alpha = \|\operatorname{tr}_{R_X}[T_{X'R_X}]\|_\infty = \|\operatorname{tr}_{R_X}\tilde\rho_{R_X}^{-1/2}\tilde\rho_{X'R_X}\tilde\rho_{R_X}^{-1/2}\|_\infty$ ensures that (C.4b) is satisfied, and hence

$$
\hat D_{X\to X'}(\tilde\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) \geqslant -\log\big\|\operatorname{tr}_{R_X}\tilde\rho_{R_X}^{-1/2}\tilde\rho_{X'R_X}\tilde\rho_{R_X}^{-1/2}\big\|_\infty\,. \qquad \text{(C.86)}
$$

Now let $\omega_{X'}\geqslant 0$ with $\operatorname{tr}\omega_{X'}=1$ such that $\operatorname{tr}\big[\omega_{X'}\cdot\operatorname{tr}_R(\tilde\rho_{R_X}^{-1/2}\tilde\rho_{X'R_X}\tilde\rho_{R_X}^{-1/2})\big] = \big\|\operatorname{tr}_{R_X}\tilde\rho_{R_X}^{-1/2}\tilde\rho_{X'R_X}\tilde\rho_{R_X}^{-1/2}\big\|_\infty$, and note that condition (C.5a) is satisfied. Now let $X_{R_X}=0$ and $Z_{X'R_X}=\tilde\rho_{R_X}^{-1}\otimes\omega_{X'}$, and we see that

$$
\tilde\rho_{R_X}^{1/2}Z_{X'R_X}\tilde\rho_{R_X}^{1/2} = \Pi_{R_X}^{\tilde\rho_{R_X}}\otimes\omega_{X'}\leqslant\mathbb{1}_{R_X}\otimes\omega_{X'}\,. \qquad \text{(C.87)}
$$

The attained value is

$$
\begin{aligned}
\operatorname{tr}\big[Z_{X'R_X}\tilde\rho_{X'R_X}\big] &= \operatorname{tr}\big[\tilde\rho_{R_X}^{-1}\otimes\omega_{X'}\cdot\tilde\rho_{X'R_X}\big]\\
&= \operatorname{tr}\big[\omega_{X'}\cdot\operatorname{tr}_{R_X}(\tilde\rho_{R_X}^{-1/2}\tilde\rho_{X'R_X}\tilde\rho_{R_X}^{-1/2})\big]\\
&= \big\|\operatorname{tr}_{R_X}\tilde\rho_{R_X}^{-1/2}\tilde\rho_{X'R_X}\tilde\rho_{R_X}^{-1/2}\big\|_\infty\,,
\end{aligned}
$$

providing us with the opposite bound to (C.86), and hence proving that

$$
\hat D_{X\to X'}(\tilde\rho_{X'R_X}\,\|\,\mathbb{1}_X,\mathbb{1}_{X'}) = -\log\big\|\operatorname{tr}_R\tilde\rho_R^{-1/2}\tilde\rho_{X'R}\tilde\rho_R^{-1/2}\big\|_\infty\,. \qquad \text{(C.88)}
$$

We now use this expression to show that

$$
\hat D_{X\to X'}(\tilde\rho_{X'R_X}\,\|\,\mathbb{1}_X,\mathbb{1}_{X'}) = -H_{\max,0}(E\,|\,X')_{\tilde\rho} = H_{\min,0}(E\,|\,R_X)_{\tilde\rho}\,. \qquad \text{(C.89)}
$$

Consider the bipartition $EX':R$ of the pure state $|\tilde\rho\rangle_{EX'R}$, and write the Schmidt decomposition $|\tilde\rho\rangle_{EX'R_X} = \tilde\rho_{EX'}^{1/2}|\Phi^{\tilde\rho}\rangle_{EX':R_X} = \tilde\rho_{R_X}^{1/2}|\Phi^{\tilde\rho}\rangle_{EX':R_X}$, with $\operatorname{tr}_{R_X}\Phi^{\tilde\rho}_{EX':R_X} = \Pi_{EX'}^{\tilde\rho_{EX'}}$. Then

$$
\begin{aligned}
\text{(C.88)} &= -\log\big\|\operatorname{tr}_{ER_X}\tilde\rho_{R_X}^{-1/2}\tilde\rho_{EX'R_X}\tilde\rho_{R_X}^{-1/2}\big\|_\infty\\
&= -\log\big\|\operatorname{tr}_{ER_X}|\Phi\rangle\langle\Phi|^{\tilde\rho}_{EX'R_X}\big\|_\infty\\
&= -\log\big\|\operatorname{tr}_E\Pi_{EX'}^{\tilde\rho_{EX'}}\big\|_\infty\\
&= -H_{\max,0}(E\,|\,X')_{\tilde\rho}\,.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\text{(C.88)} &= -\log\big\|\operatorname{tr}_{ER_X}(\tilde\rho_{R_X}^{-1/2}\tilde\rho_{EX'R_X}\tilde\rho_{R_X}^{-1/2})\big\|_\infty\\
&= -\log\big\|\operatorname{tr}_{X'}(\tilde\rho_{R_X}^{-1/2}\tilde\rho_{EX'R_X}\tilde\rho_{R_X}^{-1/2})\big\|_\infty\\
&= -\log\big\|\tilde\rho_{R_X}^{-1/2}\tilde\rho_{ER_X}\tilde\rho_{R_X}^{-1/2}\big\|_\infty = H_{\min,0}(E\,|\,R_X)_{\tilde\rho}\,,
\end{aligned}
$$

where the second equality holds because the argument of the partial trace is pure, and hence has the same spectrum on ER as on X' (by Schmidt decomposition).

We now see that

$$
\begin{aligned}
&\bar D^\epsilon_{X\to X'}(\rho_{X'R_X}\,\|\,\mathbb{1}_X,\mathbb{1}_{X'})\\
&\quad= \max_{P(\tilde\rho_{X'R_X},\rho_{X'R_X})\leqslant\epsilon}\hat D_{X\to X'}(\tilde\rho_{X'R_X}\,\|\,\mathbb{1}_X,\mathbb{1}_{X'})\\
&\quad= \max_{P(|\tilde\rho\rangle_{X'R_XE},|\rho\rangle_{X'R_XE})\leqslant\epsilon}\hat D_{X\to X'}(\tilde\rho_{X'R_X}\,\|\,\mathbb{1}_X,\mathbb{1}_{X'})\\
&\quad= \max_{P(|\tilde\rho\rangle_{X'R_XE},|\rho\rangle_{X'R_XE})\leqslant\epsilon} H_{\max,0}(E\,|\,X')_{\tilde\rho}\\
&\quad= H^\epsilon_{\max,0}(E\,|\,X')_\rho\,,
\end{aligned}
$$

where the second equality holds by properties of the purified distance (Uhlmann's theorem). An analogous argument holds for $H^\epsilon_{\min,0}(E\,|\,R_X)_\rho$. ∎

The min- and max-relative entropies already have known connections to thermodynamics [25, 28, 31] in terms of work cost of erasure and work yield of formation of a state in the presence of a heat bath. These results are recovered here, in a fully information-theoretic context.

Proposition 28 (Recovering the min- and max-relative entropies)**.** *The min-relative entropy is recovered with a trivial output state:*

$$\bar{D}^{\epsilon}_{X\to\emptyset}(\rho_{R_X} \,\|\, \Gamma_X, 1) = D^{\epsilon}_{\mathrm{min},0}(\sigma_X \,\|\, \Gamma_X)\,, \tag{C.90}$$

writing $\sigma_X = t_{R_X\to X}(\rho_{R_X})$. Furthermore the max-relative entropy is recovered with a trivial input state:

$$\bar{D}^{\epsilon}_{\emptyset\to X'}(\rho_{X'} \,\|\, 1, \Gamma'_X) = -D^{\epsilon}_{\mathrm{max}}(\rho_{X'} \,\|\, \Gamma_{X'})\,. \tag{C.91}$$

Proof of Proposition 28. For any state $\hat{\rho}_{R_X}$, consider the semidefinite program given in Proposition 12 for $2^{-\hat{D}_{X\to\emptyset}(\hat{\rho}_{R_X}\,\|\,\Gamma_X,1)}$. The choice $T_{R_X} = \Pi^{\hat{\rho}_{R_X}}_{R_X}$ along with $\alpha = \mathrm{tr}\Big(\Pi^{\hat{\rho}_{R_X}}_{R_X}\Gamma_{R_X}\Big)$ is primal feasible, hence

$$2^{-\hat{D}_{X\to\emptyset}(\hat{\rho}_{R_X}\,\|\,\Gamma_X,1)} \leqslant 2^{-D_{\mathrm{min},0}(\hat{\rho}_{R_X}\,\|\,\Gamma_{R_X})}\,. \tag{C.92}$$

In the dual problem, for any $\mu > 0$ let $Z_R = \mu\Pi^{\hat{\rho}_{R_X}}_{R_X}$ and $\omega_{X'} = 1$. Let P_{R_X} be the projector onto the eigenspaces associated with the positive (or null) eigenvalues of $(\mu\hat{\rho}_{R_X} - \Gamma_{R_X})$, and let $X_{R_X} = P_{R_X}(\mu\hat{\rho}_{R_X} - \Gamma_{R_X})P_{R_X}$. Then the dual constraints (C.16a) and (C.16b) are clearly satisfied. The attained value is

$$\begin{aligned}\mathrm{tr}(Z_{R_X}\hat{\rho}_{R_X}) - \mathrm{tr}(X_R) &= \mu\,\mathrm{tr}\,\hat{\rho}_{R_X} - \mu\,\mathrm{tr}(P_{R_X}\hat{\rho}_{R_X}) + \mathrm{tr}\,(P_{R_X}\Gamma_{R_X})\\ &\geqslant \mathrm{tr}(P_{R_X}\Gamma_{R_X}) \geqslant \mathrm{tr}\Big(\Pi^{\hat{\rho}_{R_X}}_{R_X}\Gamma_R\Big) - O(1/\mu)\,,\end{aligned} \tag{C.93}$$

where we have used Lemma 41 in the last step. If we take $\mu\to\infty$ we get successive feasible dual candidates whose attained value approaches $2^{-D_{\mathrm{min},0}(\hat{\rho}_R\,\|\,\Gamma_R)}$; hence this is the optimal value of the semidefinite program. Finally, we have

$$\begin{aligned}\bar{D}^{\epsilon}_{X\to\emptyset}(\rho_{R_X}\,\|\,\Gamma_X,1) &= \max_{\hat{\rho}_{R_X}\approx_\epsilon\rho_{R_X}} \hat{D}_{X\to\emptyset}(\hat{\rho}_{R_X}\,\|\,\Gamma_X,1)\\ &= \max_{\hat{\rho}_{R_X}\approx_\epsilon\rho_{R_X}} D_{\mathrm{min},0}(\hat{\rho}_{R_X}\,\|\,\Gamma_{R_X})\,,\\ &= D^{\epsilon}_{\mathrm{min},0}(\sigma_X\,\|\,\Gamma_X)\,.\end{aligned}$$

Let's now prove equality (C.91). For any state $\hat{\rho}_{X'}$, consider the semidefinite program given in Proposition 12 for $2^{-\hat{D}_{\emptyset\to X'}(\hat{\rho}_{X'}\,\|\,1,\Gamma_{X'})}$. The choice $T_{X'} = \rho_{X'}$ and $\alpha = \big\|\Gamma^{-1/2}_{X'}\hat{\rho}_{X'}\Gamma^{-1/2}_{X'}\big\|_\infty = 2^{D_{\mathrm{max}}(\hat{\rho}_{X'}\,\|\,\Gamma_{X'})}$ clearly satisfies the primal constraints, and thus

$$2^{-\hat{D}_{\emptyset\to X'}(\hat{\rho}_{X'}\,\|\,1,\Gamma_{X'})} \leqslant 2^{D_{\mathrm{max}}(\hat{\rho}_{X'}\,\|\,\Gamma_{X'})}\,. \tag{C.94}$$

By properties of the infinity norm, there exists a $\tau_{X'} \geqslant 0$ with $\mathrm{tr}\,\tau_{X'} = 1$ such that $\big\|\Gamma^{-1/2}_{X'}\hat{\rho}_{X'}\Gamma^{-1/2}_{X'}\big\|_\infty = \mathrm{tr}\big[\tau_{X'}\cdot\Gamma^{-1/2}_{X'}\hat{\rho}_{X'}\Gamma^{-1/2}_{X'}\big]$. Let $\omega_{X'} = \Gamma^{-1/2}_{X'}\tau_{X'}\Gamma^{-1/2}_{X'}$, $Z_{X'} = \omega_{X'}$ and $X = 0$. Then the dual constraints are trivially satisfied and the attained value is

$$\mathrm{tr}[Z_{X'}\hat{\rho}_{X'}] = \mathrm{tr}\big[\Gamma^{-1/2}_{X'}\tau_{X'}\Gamma^{-1/2}_{X'}\hat{\rho}_{X'}\big] = 2^{D_{\mathrm{max}}(\hat{\rho}_{X'}\,\|\,\Gamma_{X'})}\,. \tag{C.95}$$

The primal and dual candidates achieve the same value, and hence this is the optimal solution to the semidefinite program. We then have

$$\begin{aligned}\bar{D}^{\epsilon}_{\emptyset\to X'}(\rho_{X'}\,\|\,1,\Gamma'_X) &= \max_{\hat{\rho}_{X'}\approx_\epsilon\rho_{X'}} \hat{D}_{\emptyset\to X'}(\hat{\rho}_{X'R}\,\|\,1,\Gamma'_X)\\ &= \max_{\hat{\rho}_{X'}\approx_\epsilon\rho_{X'}} -D_{\mathrm{max}}(\hat{\rho}_{X'}\,\|\,\Gamma_{X'})\\ &= -D^{\epsilon}_{\mathrm{max}}(\rho_{X'}\,\|\,\Gamma_{X'})\,.\end{aligned}$$

■

It is clear that in Proposition 28 in the case of $\epsilon = 0$, we may replace the trivial system with $\Gamma = 1$ by a nontrivial system with arbitrary Γ, as long as it is in a pure eigenstate of the Γ operator.

Corollary 29. *Let $\Gamma_X, \Gamma_{X'} \geqslant 0$. Both following statements hold:*

(a) Let $|\mathrm{f}\rangle_{X'}$ be an eigenstate of $\Gamma_{X'}$ with eigenvalue g_f, and let σ_X be any quantum state in the support of Γ_X. Then:

$$\begin{aligned}&\hat{D}_{X\to X'}(t_{X\to R_X}(\sigma_X)\otimes|\mathrm{f}\rangle\langle\mathrm{f}|_{X'}\,\|\,\Gamma_X,\Gamma_{X'})\\ &\qquad = D_{\mathrm{min},0}(\sigma_X\,\|\,\Gamma_X) + \log g_\mathrm{f}\,.\end{aligned} \tag{C.96}$$

(b) Let $|\mathrm{i}\rangle_X$ be an eigenstate of Γ_X with eigenvalue g_i, and let $\rho_{X'}$ be any quantum state in the support of $\Gamma_{X'}$. Then:

$$\begin{aligned}&\hat{D}_{X\to X'}(t_{X\to R_X}(|\mathrm{i}\rangle\langle\mathrm{i}|_X)\otimes\rho_{X'}\,\|\,\Gamma_X,\Gamma_{X'})\\ &\qquad = -\log g_\mathrm{i} - D_{\mathrm{max}}(\rho_{X'}\,\|\,\Gamma_{X'})\,.\end{aligned} \tag{C.97}$$

Proof of Corollary 29. First consider claim (a). Invoking successively Proposition 9, Proposition 7, and Proposition 6, we have (writing $\sigma_{R_X} = t_{X\to R_X}(\sigma_X)$):

$$\begin{aligned}&\hat{D}_{X\to X'}(\sigma_{R_X}\otimes|\mathrm{f}\rangle\langle\mathrm{f}|_{X'}\,\|\,\Gamma_X,\Gamma_{X'})\\ &\quad= \hat{D}_{X\to X'}(\sigma_{R_X}\otimes|\mathrm{f}\rangle\langle\mathrm{f}|_{X'}\,\|\,\Gamma_X,\,g_\mathrm{f}|\mathrm{f}\rangle\langle\mathrm{f}|_{X'})\\ &\quad= \hat{D}_{X\to X'}(\sigma_{R_X}\otimes|\mathrm{f}\rangle\langle\mathrm{f}|_{X'}\,\|\,\Gamma_X,\,|\mathrm{f}\rangle\langle\mathrm{f}|_{X'}) + \log g_\mathrm{f}\\ &\quad= \hat{D}_{X\to\emptyset}(\sigma_{R_X}\,\|\,\Gamma_X,\,1) + \log g_\mathrm{f}\,,\end{aligned} \tag{C.98}$$

at which point we may apply Proposition 28. Claim (b) follows analogously. ■

Finally, we will see that the usual quantum relative entropy can also be recovered in the regime where we consider states of the form $\rho^{\otimes n}_{X'^nR^n}$ for $n\to\infty$. We defer this case to Section C 8, as the proof of this property requires some additional bounds we have yet to present.

7. Bounds on the coherent relative entropy

At this point, we further characterize the coherent relative entropy with bounds in terms of simpler quantities depending only on the input and output states. The main goal of this section is to prove Proposition 32 and Proposition 35, which will allow us to understand our quantity's asymptotic behavior in the i.i.d. regime.

We begin with a few upper bounds on the coherent relative entropy, given in terms of a difference of relative entropies.

Proposition 30. *We have the upper bound*

$$\hat{D}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) \leqslant D(\sigma_X\,\|\,\Gamma_X) - D(\rho_{X'}\,\|\,\Gamma_{X'})\,, \tag{C.99}$$

writing $\sigma_X = t_{R_X\to X}(\rho_{R_X})$

Proof of Proposition 30. Consider the optimal solution $T_{X'R_X}$ and α to the primal semidefinite program of Proposition 12, and let $\mathcal{T}_{X\to X'}$ be the completely positive map corresponding to $T_{X'R_X}$, i.e. defined by $\mathcal{T}_{X\to X'}(\cdot) = \mathrm{tr}_{R_X}[T_{X'R_X}\,t_{X\to R_X}(\cdot)]$. The mapping defined in this way is completely positive since $T_{X'R_X}\geqslant 0$ and is trace-nonincreasing thanks to condition (C.4a).

The map $\mathcal{T}_{X\to X'}$ thus satisfies the conditions of item (i) of Proposition 3. Hence, invoking item (ii) of that proposition, let $\tilde{\Phi}_{XA\to X'A'}$ be a trace non-increasing Γ-sub-preserving map for large enough A, A', with $\Gamma_A = \mathbb{1}_A$, $\Gamma_{A'} = \mathbb{1}_{A'}$, satisfying

$$\tilde{\Phi}_{XA\to X'A'}\Big(\sigma_{XR_X} \otimes \Big(2^{-\lambda_1}\mathbb{1}_{2^{\lambda_1}}\Big)\Big) = \rho_{X'R_X} \otimes \Big(2^{-\lambda_2}\mathbb{1}_{2^{\lambda_2}}\Big)\,, \tag{C.100}$$

with $\alpha = 2^{-(\lambda_1-\lambda_2)}$ and $|\sigma\rangle_{XR_X} = \rho_{R_X}^{1/2}\,|\Phi\rangle_{X:R_X}$. (If α is irrational, the following argument may be applied to arbitrary good rational approximations to α.)

Now, dilate $\tilde{\Phi}_{XA\to X'A'}$ using Proposition 2 to a trace-preserving, Γ-preserving map $\Phi_{XAX'A'Q\to XAX'A'Q}$ with states $|\mathrm{x}\rangle_X$, $|\mathrm{a}\rangle_A$, $|\mathrm{i}\rangle_Q$, $|\mathrm{x}'\rangle_{X'}$, $|\mathrm{a}'\rangle_{A'}$, $|\mathrm{f}\rangle_Q$ (all of them being eigenstates of the respective Γ operators), satisfying

$$\Phi_{XAX'A'Q}(\Gamma_{XAX'A'Q}) = \Gamma_{XAX'A'Q}\,; \tag{C.101a}$$

$$\begin{aligned}\Phi_{XAX'A'Q}\Big(\sigma_{XR_X} \otimes \Big(2^{-\lambda_1}\mathbb{1}^{A}_{2^{\lambda_1}}\Big) \otimes |\mathrm{x'a'i}\rangle\langle\mathrm{x'a'i}|_{X'A'Q}\Big)\\ = \rho_{X'R_X} \otimes \Big(2^{-\lambda_2}\mathbb{1}^{A'}_{2^{\lambda_2}}\Big) \otimes |\mathrm{xaf}\rangle\langle\mathrm{xaf}|_{XAQ}\,;\quad \text{and}\end{aligned} \tag{C.101b}$$

$$\langle\mathrm{x\,a\,f}|\Gamma_{XAQ}|\mathrm{x\,a\,f}\rangle_{XAQ} = \langle\mathrm{x'a'i}|\Gamma_{X'A'Q}|\mathrm{x'a'i}\rangle_{X'A'Q}\,. \tag{C.101c}$$

Using Proposition 16 recalling that $\Gamma_A = \mathbb{1}_A$, we see that

$$D\Big(2^{-\lambda_1}\mathbb{1}^{A}_{2^{\lambda_1}}\,\Big\|\,\Gamma_A\Big) = -\log\operatorname{tr}\Big(\mathbb{1}^{A}_{2^{\lambda_1}}\,\Gamma_A\Big) = -\lambda_1\,; \tag{C.102a}$$

$$D\Big(2^{-\lambda_2}\mathbb{1}^{A'}_{2^{\lambda_2}}\,\Big\|\,\Gamma_{A'}\Big) = -\log\operatorname{tr}\Big(\mathbb{1}^{A}_{2^{\lambda_2}}\,\Gamma_A\Big) = -\lambda_2\,, \tag{C.102b}$$

as well as for any pure eigenstate y of any positive semidefinite Γ,

$$D(|\mathrm{y}\rangle\langle\mathrm{y}|\,\|\,\Gamma) = -\log\operatorname{tr}\langle\mathrm{y}|\Gamma|\mathrm{y}\rangle\,. \tag{C.102c}$$

Then, by the data processing inequality for the relative entropy and with (C.101b),

$$\begin{aligned}0 &\leqslant D\Big(\sigma_X \otimes (2^{-\lambda_1}\mathbb{1}^{A}_{2^{\lambda_1}}) \otimes |\mathrm{x'a'i}\rangle\langle\mathrm{x'a'i}|_{X'A'Q}\,\Big\|\,\Gamma_{XAX'A'Q}\Big)\\ &\quad - D\Big(\rho_{X'} \otimes (2^{-\lambda_2}\mathbb{1}^{A'}_{2^{\lambda_2}}) \otimes |\mathrm{xaf}\rangle\langle\mathrm{xaf}|_{XAQ}\,\Big\|\,\Gamma_{XAX'A'Q}\Big)\\ &= D(\sigma_X\,\|\,\Gamma_X) + D(2^{-\lambda_1}\mathbb{1}^{A}_{2^{\lambda_1}}\,\|\,\Gamma_A) + D(|\mathrm{x'a'i}\rangle\langle\mathrm{x'a'i}|_{X'A'Q}\,\|\,\Gamma_{X'A'Q})\\ &\quad - D(\rho_{X'}\,\|\,\Gamma_{X'}) - D(2^{-\lambda_2}\mathbb{1}^{A'}_{2^{\lambda_2}}\,\|\,\Gamma_{A'}) - D(|\mathrm{x\,a\,f}\rangle\langle\mathrm{x\,a\,f}|_{XAQ}\,\|\,\Gamma_{XAQ})\\ &= D(\sigma_X\,\|\,\Gamma_X) - D(\rho_{X'}\,\|\,\Gamma_{X'}) - \lambda_1 + \lambda_2\\ &\quad - \log\langle\mathrm{x'a'i}|\Gamma_{X'A'Q}|\mathrm{x'a'i}\rangle + \log\langle\mathrm{x\,a\,f}|\Gamma_{XAQ}|\mathrm{x\,a\,f}\rangle\\ &= D(\sigma_X\,\|\,\Gamma_X) - D(\rho_{X'}\,\|\,\Gamma_{X'}) - \lambda_1 + \lambda_2\,,\end{aligned} \tag{C.103}$$

where we invoked the condition (C.101c) in the last step. We then have

$$\hat{D}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X, \Gamma_{X'}) = \lambda_1 - \lambda_2 \leqslant D(\sigma_X\,\|\,\Gamma_X) - D(\rho_{X'}\,\|\,\Gamma_{X'})\,.$$

■

The following upper bound is easy to prove, although it has not found tremendous use.

Proposition 31. *The coherent relative entropy may be upper bounded as*

$$\begin{aligned}&\hat{D}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})\\ &\qquad \leqslant D_{\max}(\sigma_X\,\|\,\Gamma_X) - D_{\max}(\rho_{X'}\,\|\,\Gamma_{X'})\,,\end{aligned} \tag{C.104}$$

writing $\sigma_X = t_{X\to R_X}(\rho_{R_X})$

Proof of Proposition 31. Consider an optimal solution $T_{X'R_X}$ and α for the primal semidefinite program. Then we have via the semidefinite constraints

$$\begin{aligned}\rho_{X'} = \operatorname{tr}_{R_X}\big[T_{X'R_X}\rho_{R_X}\big] &\leqslant 2^{D_{\max}(\rho_{R_X}\,\|\,\Gamma_{R_X})}\operatorname{tr}_{R_X}\big[T_{X'R_X}\Gamma_{R_X}\big]\\ &\leqslant \alpha\,2^{D_{\max}(\rho_{R_X}\,\|\,\Gamma_{R_X})}\,\Gamma_{X'}\,.\end{aligned} \tag{C.105}$$

By definition, we have

$$2^{D_{\max}(\rho_{X'}\,\|\,\Gamma_{X'})} = \min\{\mu:\ \mu\,\Gamma_{X'} \geqslant \rho_{X'}\}\,, \tag{C.106}$$

and thus we see that $\alpha 2^{D_{\max}(\rho_{R_X}\,\|\,\Gamma_{R_X})}$ is a candidate μ in this minimization. Hence $2^{D_{\max}(\rho_{X'}\,\|\,\Gamma_{X'})} \leqslant \alpha\,2^{D_{\max}(\rho_{R_X}\,\|\,\Gamma_{R_X})}$ and

$$\alpha \geqslant 2^{D_{\max}(\rho_{X'}\,\|\,\Gamma_{X'}) - D_{\max}(\rho_{R_X}\,\|\,\Gamma_{R_X})}\,. \tag{C.107}$$

The claim then follows from $\hat{D}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) = -\log\alpha$. ■

The last of the upper bounds holds for the smooth coherent relative entropy. The present upper bound will be used to prove one direction of the asymptotic equipartition property.

Proposition 32. *Let $\rho_{X'R_X}$ be any quantum state, and denote the corresponding input state by $\sigma_X = t_{R_X\to X}(\rho_{R_X})$. Then for any $\epsilon, \epsilon', \epsilon'' \geqslant 0$ such that $\bar{\epsilon} := \epsilon + \epsilon' + 2\epsilon'' < 1$,*

$$\begin{aligned}&\bar{D}^{\epsilon''}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})\\ &\leqslant D^{\epsilon}_{\max}(\sigma_X\,\|\,\Gamma_X) - D^{\epsilon'}_{\min,0}(\rho_{X'}\,\|\,\Gamma_{X'}) - \log(1-\bar{\epsilon})\,.\end{aligned} \tag{C.108}$$

Proof of Proposition 32. Let $\bar{\rho}_{X'R_X}$ be the quantum state which achieves the optimum for $\bar{D}^{\epsilon''}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})$, i.e., satisfying $P(\bar{\rho}_{X'R},\rho_{X'R}) \leqslant \epsilon''$ and $\bar{D}^{\epsilon''}_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'}) = \bar{D}_{X\to X'}(\bar{\rho}_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})$. The proof proceeds by constructing dual candidates for $2^{-\bar{D}_{X\to X'}(\bar{\rho}_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})}$ in (C.14) achieving the value in the claim. Define the quantum states $\tilde{\sigma}_X$, $\tilde{\rho}_{X'}$ as the optimal ones in the optimizations defining the smooth min and max relative entropies, i.e., satisfying $P(\tilde{\sigma}_X,\sigma_X) \leqslant \epsilon$, $P(\tilde{\rho}_{X'},\rho_{X'}) \leqslant \epsilon'$, as well as

$$D^{\epsilon}_{\max}(\sigma_X\,\|\,\Gamma_X) = D_{\max}(\tilde{\sigma}_X\,\|\,\Gamma_X)\,; \tag{C.109a}$$

$$D^{\epsilon'}_{\min,0}(\rho_{X'}\,\|\,\Gamma_{X'}) = D_{\min,0}(\tilde{\rho}_{X'}\,\|\,\Gamma_{X'})\,. \tag{C.109b}$$

Let

$$\mu = 2^{-D_{\max}(\tilde{\sigma}_X\,\|\,\Gamma_X)}\left(\operatorname{tr}\Pi^{\tilde{\rho}_{X'}}_{X'}\,\Gamma_{X'}\right)^{-1}\,; \tag{C.110a}$$

$$Z_{X'R_X} = \mu\,\Pi^{\tilde{\rho}_{X'}}_{X'} \otimes \mathbb{1}_{R_X}\,; \tag{C.110b}$$

$$\omega_{X'} = \left[\operatorname{tr}(\Pi^{\tilde{\rho}_{X'}}_{X'}\,\Gamma_{X'})\right]^{-1}\Pi^{\tilde{\rho}_{X'}}_{X'}\,. \tag{C.110c}$$

Condition (C.16a) is automatically satisfied. Writing $\tilde{\sigma}_{R_X} = t_{X\to R_X}(\tilde{\sigma}_X)$, we have $D(\bar{\rho}_{R_X},\tilde{\sigma}_{R_X}) \leqslant P(\bar{\rho}_{R_X},\tilde{\sigma}_{R_X}) \leqslant P(\bar{\rho}_{R_X},\rho_{R_X}) + P(\rho_{R_X},\tilde{\sigma}_{R_X}) \leqslant \epsilon'' + \epsilon$; hence, there exists $\Delta_{R_X} \geqslant 0$ such that $\bar{\rho}_{R_X} \leqslant \tilde{\sigma}_{R_X} + \Delta_{R_X}$ with $\operatorname{tr}\Delta_{R_X} \leqslant \epsilon'' + \epsilon$. Then,

$$\begin{aligned}\bar{\rho}^{1/2}_{R_X}\,Z_{X'R_X}\,\bar{\rho}^{1/2}_{R_X} &= \mu\,\Pi^{\tilde{\rho}_{X'}}_{X'} \otimes \bar{\rho}_{R_X}\\ &\leqslant \mu\,\Pi^{\tilde{\rho}_{X'}}_{X'} \otimes (\tilde{\sigma}_{R_X} + \Delta_{R_X})\\ &\leqslant \mu\,\Pi^{\tilde{\rho}_{X'}}_{X'} \otimes (2^{D_{\max}(\tilde{\sigma}_{R_X}\,\|\,\Gamma_{R_X})}\,\Gamma_{R_X} + \Delta_{R_X})\\ &\leqslant \omega_{X'} \otimes \Gamma_{R_X} + \mu\,\mathbb{1}_{X'} \otimes \Delta_{R_X}\,,\end{aligned} \tag{C.111}$$

and we may define $X_{R_X} = \mu\,\Delta_R$ in order for constraint (C.16b) to be also satisfied. The attained dual objective value is

$$\text{obj.} = \operatorname{tr}(Z_{X'R_X}\,\bar{\rho}_{X'R_X}) - \operatorname{tr}(X_{R_X}) = \mu\left(\operatorname{tr}(\Pi^{\tilde{\rho}_{X'}}_{X'}\,\bar{\rho}_{X'}) - \epsilon'' - \epsilon\right)\,. \tag{C.112}$$

Analogously to the input state, now we have for the output state $D(\bar{\rho}_{X'},\tilde{\rho}_{X'}) \leqslant P(\bar{\rho}_{X'},\tilde{\rho}_{X'}) \leqslant P(\bar{\rho}_{X'},\rho_{X'}) + P(\rho_{X'},\tilde{\rho}_{X'}) \leqslant \epsilon'' + \epsilon'$; there must exist $\Delta_{X'} \geqslant 0$ with $\bar{\rho}_{X'} \geqslant \tilde{\rho}_{X'} - \Delta_{X'}$ and $\operatorname{tr}\Delta_{X'} \leqslant \epsilon'' + \epsilon'$. Hence, $\operatorname{tr}(\Pi^{\tilde{\rho}_{X'}}_{X'}\,\bar{\rho}_{X'}) \geqslant \operatorname{tr}(\Pi^{\tilde{\rho}_{X'}}_{X'}\,\tilde{\rho}_{X'}) - \operatorname{tr}(\Pi^{\tilde{\rho}_{X'}}_{X'}\,\Delta_{X'}) \geqslant 1 - \epsilon'' - \epsilon'$. Thus,

$$\text{(C.112)} \geqslant \mu\,(1 - \epsilon - \epsilon' - 2\epsilon'')\,. \tag{C.113}$$

The claim follows by noting that $-\log\mu = D^{\epsilon}_{\max}(\sigma_X\,\|\,\Gamma_X) - D^{\epsilon'}_{\min,0}(\rho_{X'}\,\|\,\Gamma_{X'})$.

■

In order to formulate lower bounds on the coherent relative entropy, we introduce a generalization of the *Rob entropy* or *smooth S-entropy* [107]:

$$D_{\mathrm{r}}(\rho \,\|\, \Gamma) = -\log \left\|\rho^{-1/2}\Gamma\rho^{-1/2}\right\|_\infty \\ = -\log\, \min\{\nu : \ \nu\rho \geqslant \Pi^\rho\Gamma\Pi^\rho\}\ ; \tag{C.114}$$

$$D^\epsilon_{\mathrm{r}}(\rho \,\|\, \Gamma) = \max_{\hat\rho\approx_\epsilon\rho} D_{\mathrm{r}}(\hat\rho \,\|\, \Gamma)\ . \tag{C.115}$$

Proposition 33. *We have the lower bound*

$$\hat D_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \geqslant D_{\mathrm{r}}(\sigma_X \,\|\, \Gamma_X) - D_{\max}(\rho_{X'} \,\|\, \Gamma_{X'})\ , \tag{C.116}$$

with $\sigma_X = t_{X\to R_X}(\rho_{R_X})$.

Proof of Proposition 33. Choose the primal candidate $T_{X'R_X} = \rho_{R_X}^{-1/2}\rho_{X'R_X}\rho_{R_X}^{-1/2}$. We have $\operatorname{tr}_{X'} T_{X'R_X} = \rho_{R_X}^{-1/2}\rho_{R_X}\rho_{R_X}^{-1/2} = \Pi^{\rho_{R_X}}_{R_X} \leqslant \mathbb{1}_{R_X}$ so our candidate satisifes (C.15a). Also (C.15c) is satisfied by construction, and $\operatorname{tr}_{R_X}(T_{X'R_X}\Gamma_{R_X})$ is in the support of $\rho_{X'}$ and hence it lies in the support of $\Gamma_{X'}$. According to Proposition 11 we choose $\alpha = \|\Gamma_{X'}^{-1/2}\operatorname{tr}_{R_X}[T_{X'R_X}\Gamma_{R_X}]\Gamma_{X'}^{-1/2}\|_\infty$ and

$$\begin{aligned} &2^{-\hat D_{X\to X'}(\rho_{X'R_X}\,\|\,\Gamma_X,\Gamma_{X'})} \\ &\quad\leqslant \alpha = \|\Gamma_{X'}^{-1/2}\operatorname{tr}_{R_X}[T_{X'R_X}\Gamma_{R_X}]\Gamma_{X'}^{-1/2}\|_\infty \\ &\quad= \|\Gamma_{X'}^{-1/2}\operatorname{tr}_{R_X}\Big[T_{X'R_X}\Pi^{\rho_{R_X}}_{R_X}\Gamma_{R_X}\Pi^{\rho_{R_X}}_{R_X}\Big]\Gamma_{X'}^{-1/2}\|_\infty \\ &\quad\leqslant 2^{-D_{\mathrm{r}}(\rho_{R_X}\,\|\,\Gamma_{R_X})}\|\Gamma_{X'}^{-1/2}\operatorname{tr}_{R_X}[T_{X'R_X}\rho_{R_X}]\Gamma_{X'}^{-1/2}\|_\infty\ , \end{aligned} \tag{C.117}$$

since by definition $\rho_{R_X}^{-1/2}\Gamma_{R_X}\rho_{R_X}^{-1/2} \leqslant 2^{-D_{\mathrm{r}}(\rho_{R_X}\,\|\,\Gamma_{R_X})}\mathbb{1}$ and thus $\Pi^{\rho_{R_X}}_{R_X}\Gamma_{R_X}\Pi^{\rho_{R_X}}_{R_X} \leqslant 2^{-D_{\mathrm{r}}(\rho_{R_X}\,\|\,\Gamma_{R_X})}\rho_{R_X}$. Then

$$\begin{aligned} \text{(C.117)} &= 2^{-D_{\mathrm{r}}(\rho_{R_X}\,\|\,\Gamma_{R_X})}\|\Gamma_{X'}^{-1/2}\rho_{X'}\Gamma_{X'}^{-1/2}\|_\infty \\ &= 2^{-D_{\mathrm{r}}(\sigma_X\,\|\,\Gamma_X)}\,2^{D_{\max}(\rho_{X'}\,\|\,\Gamma_{X'})}\ . \end{aligned}$$ ∎

The quantity $D_{\mathrm{r}}(\cdot \,\|\, \cdot)$, when smoothed, is essentially equal to the min-relative entropy: These two differ by a term which is logarithmic in the failure probability. In this way, the smooth quantity $D^\epsilon_{\mathrm{r}}(\cdot \,\|\, \cdot)$ may be related to a better known quantity with an operational interpretation.

Proposition 34. *Let* $\epsilon > 0$. *Then*

$$D^\epsilon_{\mathrm{r}}(\rho \,\|\, \Gamma) \geqslant D_{\min,0}(\rho \,\|\, \Gamma) + \log\epsilon'\ , \tag{C.118}$$

where $\epsilon' = \epsilon^2/(2+\epsilon^2)$, *or equivalently,* $\epsilon = \sqrt{2\epsilon'/(1-\epsilon')}$.

Proof of Proposition 34. The proof of this proposition proceeds via the hypothesis testing relative entropy, $D^\eta_{\mathrm{H}}(\rho \,\|\, \Gamma)$. Let $\epsilon' = \epsilon^2/(2+\epsilon^2)$ and let $\eta = 1-\epsilon'$. The hypothesis testing relative entropy can be written as the solution of a semidefinite program [106]. Specifically, there exists $Q \geqslant 0$, $\mu \geqslant 0$ and $X \geqslant 0$ such that

$$2^{-D^\eta_{\mathrm{H}}(\rho\,\|\,\Gamma)} = \frac{1}{\eta}\operatorname{tr}[Q\Gamma] = \mu - \frac{\operatorname{tr} X}{\eta}\ , \tag{C.119}$$

with Q, μ and X satisfying the conditions

$$Q \leqslant \mathbb{1}\ ; \tag{C.120a}$$
$$\operatorname{tr}[Q\rho] \geqslant \eta\ ; \tag{C.120b}$$
$$\mu\rho \leqslant \Gamma + X\ . \tag{C.120c}$$

In addition, the complementary slackness relations for these variables read

$$XQ = X\ ; \tag{C.121a}$$
$$\operatorname{tr}(Q\rho) = \eta\ ; \tag{C.121b}$$
$$Q(\mu\rho - \Gamma - X) = 0\ . \tag{C.121c}$$

Define $\bar\rho = \Pi^Q\rho\Pi^Q$, where Π^Q is the projector onto the support of Q. Apply $Q^{-1}(\cdot)\Pi^Q$ onto (C.121c) to obtain

$$\mu\bar\rho = \Pi^Q\Gamma\Pi^Q + \Pi^Q X\Pi^Q \geqslant \Pi^Q\Gamma\Pi^Q\ . \tag{C.122}$$

In addition, because $\Pi^Q\Gamma\Pi^Q$ has support on Π^Q, then $\bar\rho$ must also have support on the full of Π^Q, i.e. $\Pi^{\bar\rho} = \Pi^Q$. So, by definition of $D_{\mathrm{r}}(\bar\rho \,\|\, \Gamma)$ have that

$$2^{-D_{\mathrm{r}}(\bar\rho\,\|\,\Gamma)} \leqslant \mu\ . \tag{C.123}$$

Also, define $\bar\rho' = \bar\rho/\operatorname{tr}\bar\rho$, and we can see by Lemma 43 that $P(\rho,\bar\rho') \leqslant \sqrt{2\epsilon'/(1-\epsilon')} = \epsilon$. Also, $2^{-D_{\mathrm{r}}(\bar\rho'\,\|\,\Gamma)} \leqslant 2^{-D_{\mathrm{r}}(\bar\rho\,\|\,\Gamma)}$ by definition of $D_{\mathrm{r}}(\cdot \,\|\, \cdot)$. Then $\bar\rho'$ is a valid optimization candidate in the definition of $D^{\bar\epsilon}_{\mathrm{r}}(\rho \,\|\, \Gamma)$ and

$$2^{-D^{\bar\epsilon}_{\mathrm{r}}(\rho\,\|\,\Gamma)} \leqslant 2^{-D_{\mathrm{r}}(\bar\rho'\,\|\,\Gamma)} \leqslant \mu\ . \tag{C.124}$$

It thus remains to show that $\mu \leqslant \epsilon'^{-1}\,2^{-D_{\min,0}(\rho\,\|\,\Gamma)}$. Apply $\operatorname{tr}(\Pi^\rho(\cdot))$ onto the constraint (C.120c) to obtain

$$\mu \leqslant \operatorname{tr}(\Pi^\rho\Gamma) + \operatorname{tr}(\Pi^\rho X) \leqslant \operatorname{tr}(\Pi^\rho\Gamma) + \operatorname{tr}(X)\ . \tag{C.125}$$

Now, because of (C.119), we have $0 \leqslant \operatorname{tr}[Q\Gamma] = \mu\eta - \operatorname{tr} X$, and thus $\operatorname{tr} X \leqslant \mu\eta$. Combining with (C.125) gives

$$\mu(1-\eta) \leqslant \operatorname{tr}(\Pi^\rho\Gamma)\ ; \tag{C.126}$$

since $\epsilon' = 1-\eta$ and $\operatorname{tr}(\Pi^\rho\Gamma) = 2^{-D_{\min,0}(\rho\,\|\,\Gamma)}$ we have $\mu \leqslant (1/\epsilon')\,2^{-D_{\min,0}(\rho\,\|\,\Gamma)}$ and the claim follows. ∎

The following proposition gives a lower bound to the smooth coherent relative entropy. This will prove crucial to the proof of the asymptotic equipartition theorem.

Proposition 35. *Let* $\epsilon', \epsilon'' \geqslant 0$ *and* $\epsilon''' > 0$. *Let* $\epsilon \geqslant 2\sqrt{2\epsilon'} + 2\sqrt{2(\epsilon''+\epsilon''')}$. *Then*

$$\begin{aligned} &\bar D^\epsilon_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \\ &\qquad\geqslant D^{\epsilon''}_{\min,0}(\sigma_X \,\|\, \Gamma_X) - D^{\epsilon'}_{\max}(\rho_{X'} \,\|\, \Gamma_{X'}) + \log\frac{\epsilon'''^2}{2+\epsilon'''^2}\ , \end{aligned} \tag{C.127}$$

where $\sigma_X = t_{X\to R_X}(\rho_{R_X})$.

Proof of Proposition 35. Let $\tilde\rho_{R_X}$, $\tilde\rho_{X'}$ be quantum states which are optimal smoothed states for the quantities

$$D^{\epsilon''}_{\min,0}(\rho_{R_X} \,\|\, \Gamma_{R_X}) = D_{\min,0}(\tilde\rho_{R_X} \,\|\, \Gamma_{R_X})\ . \tag{C.128a}$$
$$D^{\epsilon'}_{\max}(\rho_{X'} \,\|\, \Gamma_{X'}) = D_{\max}(\tilde\rho_{X'} \,\|\, \Gamma_{X'})\ . \tag{C.128b}$$

With $\epsilon''' > 0$ and using Proposition 34, we know that

$$D^{\epsilon'''}_{\mathrm{r}}(\tilde\rho_{R_X} \,\|\, \Gamma_{R_X}) \geqslant D_{\min,0}(\tilde\rho_{R_X} \,\|\, \Gamma_{R_X}) + \log\frac{\epsilon'''^2}{2+\epsilon'''^2}\ . \tag{C.129}$$

Let $\tilde{\tilde\rho}_{R_X}$ be the optimal smoothed state for $D^{\epsilon'''}_{\mathrm{r}}(\tilde\rho_{R_X} \,\|\, \Gamma_{R_X})$, such that

$$D_{\mathrm{r}}(\tilde{\tilde\rho}_{R_X} \,\|\, \Gamma_{R_X}) = D^{\epsilon'''}_{\mathrm{r}}(\tilde\rho_{R_X} \,\|\, \Gamma_{R_X})\ . \tag{C.130}$$

At this point, we have

$$\begin{aligned} &D_{\mathrm{r}}(\tilde{\tilde\rho}_{R_X} \,\|\, \Gamma_{R_X}) - D_{\max}(\tilde\rho \,\|\, \Gamma_{X'}) \\ &\qquad\geqslant D^{\epsilon''}_{\min,0}(\rho_{R_X} \,\|\, \Gamma_{R_X}) - D^{\epsilon'}_{\max}(\rho_{X'} \,\|\, \Gamma_{X'}) + \log\frac{\epsilon'''^2}{2+\epsilon'''^2}\ , \end{aligned} \tag{C.131}$$

with

$$P(\tilde{\rho}_{X'}, \rho_{X'}) \leqslant \epsilon' \ ; \quad P(\tilde{\rho}_{R_X}, \rho_{R_X}) \leqslant \epsilon'' \ ; \quad P(\tilde{\tilde{\rho}}_{R_X}, \tilde{\rho}_{R_X}) \leqslant \epsilon''' \ . \tag{C.132}$$

Now, we'll apply Lemma 44 twice to construct a state close to $\rho_{X'R_X}$ which has marginals $\tilde{\rho}_{X'}$ and $\tilde{\tilde{\rho}}_{R_X}$ exactly. Let $\tau_{X'R_X}$ be the quantum state given by Lemma 44 satisfying

$$\tau_{X'} = \tilde{\rho}_{X'} \ ; \quad \tau_{R_X} = \rho_{R_X} \ ; \quad P(\tau_{X'R_X}, \rho_{X'R_X}) \leqslant 2\sqrt{2\epsilon'} \ . \tag{C.133}$$

Applying Lemma 44 again, let $\tau'_{X'R}$ be a quantum state close to $\tau_{X'R}$ such that

$$\tau'_{X'} = \tilde{\rho}_{X'} \ ; \quad \tau'_{R_X} = \tilde{\tilde{\rho}}_{R_X} \ ; \quad P(\tau'_{X'R_X}, \tau_{X'R_X}) \leqslant 2\sqrt{2(\epsilon'' + \epsilon''')} \ . \tag{C.134}$$

We thus have by triangle inequality

$$P(\tau'_{X'R_X}, \rho_{X'R_X}) \leqslant 2\sqrt{2\epsilon'} + 2\sqrt{2(\epsilon'' + \epsilon''')} \ . \tag{C.135}$$

By Proposition 33 we can now write

$$\begin{aligned}\hat{D}_{X\to X'}(\tau'_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) &\geqslant D_{\mathrm{r}}(\tau'_{R_X} \,\|\, \Gamma_{R_X}) - D_{\max}(\tau'_{X'} \,\|\, \Gamma_{X'}) \\ &= D_{\mathrm{r}}(\tilde{\tilde{\rho}}_{R_X} \,\|\, \Gamma_{R_X}) - D_{\max}(\tilde{\rho}_{X'} \,\|\, \Gamma_{X'}) \ . \end{aligned} \tag{C.136}$$

Observe now that $\tau'_{X'R_X}$ is a valid optimization candidate for $\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'})$. Hence

$$\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \geqslant \hat{D}_{X\to X'}(\tau'_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \ . \tag{C.137}$$

Finally, inequality (C.137) followed by (C.136) and (C.131) provides us the seeked lower bound. ∎

We also have a bound which applies to product states, given in terms of min- and max-relative entropies of input and output. Physically, it asserts that a possible strategy for implementing the product state process matrix is to completely erase the input state (at a cost given by the min-relative entropy), and subsequently prepare the required output state (at a yield given by the max-relative entropy).

Proposition 36 (coherent relative entropy for product states)**.** *For states σ_X and $\rho_{X'}$, we have*

$$\begin{aligned}&\hat{D}_{X\to X'}(t_{X\to R_X}(\sigma_X) \otimes \rho_{X'} \,\|\, \Gamma_X, \Gamma_{X'}) \\ &\qquad \geqslant D_{\min,0}(\sigma_X \,\|\, \Gamma_X) - D_{\max}(\rho_{X'} \,\|\, \Gamma_{X'}) \ . \end{aligned} \tag{C.138}$$

Proof of Proposition 36. Write $\sigma_{R_X} = t_{X\to R_X}(\sigma_X)$. Choose $T_{X'R_X} = \Pi^{\sigma_{R_X}}_{R_X} \otimes \rho_{X'}$. This choice trivially satisfies (C.15a). Also, $\sigma^{1/2}_{R_X} T_{X'R_X} \sigma^{1/2}_{R_X} = \sigma_{R_X} \otimes \rho_{X'}$ so (C.15c) is also satisfied. We have that $\mathrm{tr}_{R_X}\, T_{X'R_X}\Gamma_{R_X}$ lies in the support of $\Gamma_{X'}$ because $\rho_{X'}$ does so, and as per Proposition 11 the optimal value of α corresponding to this $T_{X'R_X}$ is given by

$$\begin{aligned}\alpha &= \big\|\Gamma^{-1/2}_{X'}\, \mathrm{tr}_{R_X}\big[T_{X'R_X}\Gamma_{R_X}\big]\Gamma^{-1/2}_{X'}\big\|_\infty \\ &= \big\|\Gamma^{-1/2}_{X'}\, \mathrm{tr}_{R_X}\Big[(\Pi^{\sigma_{R_X}}_{R_X} \otimes \rho_{X'})\Gamma_{R_X}\Big]\Gamma^{-1/2}_{X'}\big\|_\infty \\ &= \mathrm{tr}_{R_X}\big[\Pi^{\sigma_{R_X}}_{R_X}\Gamma_{R_X}\big]\big\|\Gamma^{-1/2}_{X'}\rho_{X'}\Gamma^{-1/2}_{X'}\big\|_\infty \\ &= 2^{-D_{\min,0}(\sigma_{R_X} \,\|\, \Gamma_{R_X})}\, 2^{D_{\max}(\rho_{X'} \,\|\, \Gamma_{X'})} \ . \end{aligned} \tag{C.139}$$

This choice of α and $T_{X'R_X}$ is feasible for $2^{-\hat{D}_{X\to X'}(\sigma_{R_X}\otimes\rho_{X'} \,\|\, \Gamma_X,\Gamma_{X'})}$, hence

$$\begin{aligned}&\hat{D}_{X\to X'}(\sigma_{R_X} \otimes \rho_{X'} \,\|\, \Gamma_X, \Gamma_{X'}) \\ &\qquad \geqslant D_{\min,0}(\sigma_X \,\|\, \Gamma_X) - D_{\max}(\rho_{X'} \,\|\, \Gamma_{X'}) \ . \end{aligned}$$

∎

8. Asymptotic equipartition property

Finally, the coherent relative entropy also obeys an asymptotic equipartition property in the i.i.d. limit. In this limit, the coherent relative entropy converges to the difference of relative entropies of the input and the output to the respective Γ operators.

Both versions of the coherent relative entropy we have introduced have the same asymptotic behavior for small ϵ. For completeness we present the detailed statements, including the ranges of ϵ for which the property is proven for each quantity.

Proposition 37 (Asymptotic equipartition property)**.** *For any $\Gamma_X, \Gamma_{X'} \geqslant 0$, for any quantum state $\rho_{X'R_X}$, and for any $0 < \epsilon < 1/2$,*

$$\begin{aligned}&\lim_{n\to\infty} \frac{1}{n} \bar{D}^{\epsilon}_{X^n\to X'^n}\big(\rho^{\otimes n}_{X'R_X} \,\big\|\, \Gamma^{\otimes n}_X, \Gamma^{\otimes n}_{X'}\big) \\ &\qquad = D(\sigma_X \,\|\, \Gamma_X) - D(\rho_{X'} \,\|\, \Gamma_{X'}) \ , \end{aligned} \tag{C.140}$$

where $\sigma_X = t_{X\to X'}(\rho_{R_X})$.

Similarly, using the same definitions and for any $0 < \epsilon < (18)^{-4}$,

$$\begin{aligned}&\lim_{n\to\infty} \frac{1}{n} \hat{D}^{\epsilon}_{X^n\to X'^n}\big(\rho^{\otimes n}_{X'R_X} \,\big\|\, \Gamma^{\otimes n}_X, \Gamma^{\otimes n}_{X'}\big) \\ &\qquad = D(\sigma_X \,\|\, \Gamma_X) - D(\rho_{X'} \,\|\, \Gamma_{X'}) \ . \end{aligned} \tag{C.141}$$

(Proof on page 30.)

The proof of the asymptotic equipartition follows from bounds we have derived using the min- and max-relative entropies. The latter have known asymptotic behavior, summarized as follows:

Proposition 38 (Asymptotic equipartition for min- and max-relative entropies [62, 106])**.** *The known asymptotic behavior of the min and max-relative entropy are, for any fixed $0 < \epsilon < 1/2$,*

$$\lim_{n\to\infty} \frac{1}{n} D^{\epsilon}_{\min,0}(\sigma^{\otimes n}_X \,\|\, \Gamma^{\otimes n}_X) = D(\sigma_X \,\|\, \Gamma_X) \ ; \tag{C.142a}$$

$$\lim_{n\to\infty} \frac{1}{n} D^{\epsilon}_{\max}(\rho^{\otimes n}_{X'} \,\|\, \Gamma^{\otimes n}_{X'}) = D(\rho_{X'} \,\|\, \Gamma_{X'}) \ . \tag{C.142b}$$

Proof of Proposition 38. Because of the peculiarities of our definitions and form of the asymptotic equipartition statement (e.g., no limit $\epsilon \to 0$) which differ slightly from the existing results in the literature, we provide a formal proof for completeness. This proof proceeds by bounding the min- and max-relative entropies from both sides by an expression involving the hypothesis testing relative entropy. Quantum Stein's lemma [108, 109] applied to the hypothesis testing relative entropy [106, Prop. 3.4] then automatically allows us to recover the quantum relative entropy in the i.i.d. limit. Additive corrective terms which do not scale with the number of copies are also automatically suppressed in this regime.

Let $\tilde{\sigma}$ be optimal for $D^{\epsilon}_{\min,0}(\sigma \,\|\, \Gamma)$, i.e., $D^{\epsilon}_{\min,0}(\sigma \,\|\, \Gamma) = D_{\min,0}(\tilde{\sigma} \,\|\, \Gamma) = -\log \mathrm{tr}[\Pi^{\tilde{\sigma}}\, \Gamma]$ with $P(\sigma, \tilde{\sigma}) \leqslant \epsilon$. As $\sigma \geqslant \tilde{\sigma} - \Delta$ for some $\Delta \geqslant 0$ with $\mathrm{tr}\,\Delta \leqslant \epsilon$, we have that $\mathrm{tr}[\Pi^{\tilde{\sigma}}\sigma] \geqslant 1 - \mathrm{tr}(\Pi^{\tilde{\sigma}}\Delta) \geqslant 1 - \epsilon$. Then $\Pi^{\tilde{\sigma}}$ is feasible in the primal program for $2^{-D^{1-\epsilon}_{\mathrm{H}}(\sigma \,\|\, \Gamma)}$, achieving the value $(1-\epsilon)^{-1}\,\mathrm{tr}(\Pi^{\tilde{\sigma}}\Gamma)$. Hence, for any $0 < \epsilon < 1$,

$$D^{\epsilon}_{\min,0}(\sigma \,\|\, \Gamma) \leqslant D^{1-\epsilon}_{\mathrm{H}}(\sigma \,\|\, \Gamma) - \log(1-\epsilon) \ . \tag{C.143}$$

Conversely, for any $0 < \epsilon' < 1/2$ to be fixed later, let Q be primal optimal for $2^{-D_{\mathrm{H}}^{1-\epsilon'}(\sigma\,\|\,\Gamma)} = (1-\epsilon')^{-1}\operatorname{tr}(Q\Gamma)$ with $\operatorname{tr}(Q\sigma) \geqslant 1-\epsilon'$. For $\eta = \epsilon'$, Let P^η be the projector onto the eigenspaces of Q associated to eigenvalues greater than or equal to η, and hence satisfying $\eta P^\eta \leqslant Q$. It follows that $\operatorname{tr}(Q\Gamma) \geqslant \eta\operatorname{tr}(P^\eta\Gamma)$. Now, define $\tilde\sigma = P^\eta\,\sigma\,P^\eta/\operatorname{tr}(P^\eta\sigma)$, noting that $\operatorname{tr}(P^\eta\sigma) \geqslant \operatorname{tr}(P^\eta Q P^\eta\sigma) \geqslant \operatorname{tr}(Q\sigma) - \operatorname{tr}((\mathbb{1}-P^\eta)\,Q\,(\mathbb{1}-P^\eta)\,\sigma) \geqslant 1-\epsilon'-\eta$ (recall that all eigenvalues of $(\mathbb{1}-P^\eta)\,Q\,(\mathbb{1}-P^\eta)$ are less than η). Using Lemma 43, we see that $P(\tilde\sigma,\sigma) \leqslant \sqrt{2(\epsilon'+\eta)}/\sqrt{1-\epsilon'-\eta} = \sqrt{4\epsilon'/(1-2\epsilon')}$. Now $\tilde\sigma$ is a valid candidate for the smoothing in $D_{\min,0}^{\sqrt{4\epsilon'/(1-2\epsilon')}}(\sigma\,\|\,\Gamma)$, and hence $D_{\min,0}^{\sqrt{4\epsilon'/(1-2\epsilon')}}(\sigma\,\|\,\Gamma) \geqslant -\log\operatorname{tr}(P^\eta\Gamma) \geqslant -\log[\eta^{-1}\operatorname{tr}(Q\Gamma)] = -\log[((1-\epsilon')/\epsilon')\,(1-\epsilon')^{-1}\operatorname{tr}(Q\Gamma)] = D_{\mathrm{H}}^{\epsilon'}(\sigma\,\|\,\Gamma) - \log[(1-\epsilon')/\epsilon']$. For any $0 < \epsilon < 1/2$, by setting $\epsilon' = \epsilon^2/(4+2\epsilon^2)$, we have $\epsilon = \sqrt{4\epsilon'/(1-\epsilon')}$. Hence, for any $0 < \epsilon < 1/2$, we have:

$$D_{\min,0}^{\epsilon}(\sigma\,\|\,\Gamma) \geqslant D_{\mathrm{H}}^{\epsilon'}(\sigma\,\|\,\Gamma) - \log\frac{1-\epsilon'}{\epsilon'}\ . \tag{C.144}$$

For the max-relative entropy, for any ρ, Γ and for any $0 < \epsilon < 1/2$, let $\tilde\rho$ be a normalized quantum state such that $D_{\max}^{\epsilon}(\rho\,\|\,\Gamma) = D_{\max}(\tilde\rho\,\|\,\Gamma)$. Let Q be primal optimal for $2^{-D_{\mathrm{H}}^{2\epsilon}(\rho\,\|\,\Gamma)} = (2\epsilon)^{-1}\operatorname{tr}(Q\Gamma)$, such that $\operatorname{tr}(Q\rho) \geqslant 2\epsilon$. But $\tilde\rho \geqslant \rho - \Delta$ for a $\Delta \geqslant 0$ with $\operatorname{tr}\Delta \leqslant \epsilon$, since $D(\tilde\rho,\rho) \leqslant \epsilon$, and thus $\operatorname{tr}(Q\tilde\rho) \geqslant 2\epsilon - \epsilon = \epsilon$. Then Q is primal feasible also for $D_{\mathrm{H}}^{\epsilon}(\tilde\rho\,\|\,\Gamma)$ and $2^{-D_{\mathrm{H}}^{\epsilon}(\tilde\rho\,\|\,\Gamma)} \leqslant \epsilon^{-1}\operatorname{tr}(Q\Gamma) = 2\cdot 2^{-D_{\mathrm{H}}^{2\epsilon}(\rho\,\|\,\Gamma)}$. Then, using [106, Prop. 4.1], $D_{\max}(\tilde\rho\,\|\,\Gamma) \geqslant D_{\mathrm{H}}^{\epsilon}(\tilde\rho\,\|\,\Gamma) \geqslant D_{\mathrm{H}}^{2\epsilon}(\rho\,\|\,\Gamma) - 1$, and hence

$$D_{\max}^{\epsilon}(\rho\,\|\,\Gamma) \geqslant D_{\mathrm{H}}^{2\epsilon}(\rho\,\|\,\Gamma) - 1\ . \tag{C.145}$$

For a lower bound on $D_{\max}^{\epsilon}$, we invoke [106, Prop. 4.1]; however the quantity called $D_{\max}^{\epsilon}$ there optimizes over subnormalized states whereas we optimize over normalized states only, so we have to work a little more. For any subnormalized state $\tilde\rho$ with $\operatorname{tr}\tilde\rho \geqslant 1-\epsilon$, we have by definition that $2^{D_{\max}(\tilde\rho\,\|\,\Gamma)} = \|\Gamma^{-1/2}\,\tilde\rho\,\Gamma^{-1/2}\|_\infty = \operatorname{tr}(\tilde\rho)\,2^{D_{\max}(\tilde\rho/\operatorname{tr}\tilde\rho\,\|\,\Gamma)} \geqslant (1-\epsilon)\cdot 2^{D_{\max}(\tilde\rho/\operatorname{tr}\tilde\rho\,\|\,\Gamma)}$, and hence

$$\begin{aligned}\min_{\substack{\tilde\rho:\ \operatorname{tr}\tilde\rho\leqslant 1\\ P(\tilde\rho,\rho)\leqslant\epsilon}} D_{\max}(\tilde\rho\,\|\,\Gamma) &\geqslant \min_{\substack{\tilde\rho:\ \operatorname{tr}\tilde\rho\leqslant 1\\ P(\tilde\rho,\rho)\leqslant\epsilon}} D_{\max}(\tilde\rho/\operatorname{tr}\tilde\rho\,\|\,\Gamma) + \log(1-\epsilon)\\ &= D_{\max}^{\epsilon}(\rho\,\|\,\Gamma) + \log(1-\epsilon)\ .\end{aligned} \tag{C.146}$$

Then, invoking [106, Prop. 4.1] for any $0 < \epsilon < 1$, and chaining with the above inequality,

$$D_{\mathrm{H}}^{\epsilon^2/2}(\rho\,\|\,\Gamma) \geqslant D_{\max}^{\epsilon}(\rho\,\|\,\Gamma) + \log(1-\epsilon)\ . \qquad ■$$

Proof of Proposition 37. We start by upper bounding the coherent relative entropy $\bar D_{X^n\to X'^n}^{\epsilon}\big(\rho_{X'R_X}^{\otimes n}\,\big\|\,\Gamma_X^{\otimes n},\Gamma_{X'}^{\otimes n}\big)$. Thanks to Proposition 32, choosing $\tilde\epsilon = \tilde\epsilon' = (1-2\epsilon)/137414920$ and with $\tilde{\tilde\epsilon} = \tilde\epsilon + \tilde\epsilon' + 2\epsilon$,

$$\begin{aligned}&\lim_{n\to\infty}\frac{1}{n}\bar D_{X^n\to X'^n}^{\epsilon}\big(\rho_{X'R_X}^{\otimes n}\,\big\|\,\Gamma_X^{\otimes n},\Gamma_{X'}^{\otimes n}\big)\\ &\quad\leqslant \lim_{n\to\infty}\frac{1}{n}\Big[D_{\max}^{\tilde\epsilon}\big(\rho_{R_X}^{\otimes n}\,\big\|\,\Gamma_{R_X}^{\otimes n}\big) - D_{\min,0}^{\tilde\epsilon'}\big(\rho_{X'}^{\otimes n}\,\big\|\,\Gamma_{X'}^{\otimes n}\big) - \log\big[\epsilon(1-\tilde{\tilde\epsilon})\big]\Big]\\ &\quad= D(\rho_{R_X}\,\|\,\Gamma_{R_X}) - D(\rho_{X'}\,\|\,\Gamma_{X'})\ .\end{aligned} \tag{C.147}$$

The lower bound is given by Proposition 35: Choosing $\hat\epsilon' = \hat\epsilon'' = \hat\epsilon''' = \epsilon^2/197334000868$,

$$\begin{aligned}&\lim_{n\to\infty}\frac{1}{n}\bar D_{X^n\to X'^n}^{\epsilon}\big(\rho_{X'R_X}^{\otimes n}\,\big\|\,\Gamma_X^{\otimes n},\Gamma_{X'}^{\otimes n}\big)\\ &\quad\geqslant \lim_{n\to\infty}\frac{1}{n}\Big[D_{\min,0}^{\hat\epsilon''}\big(\rho_{R_X}^{\otimes n}\,\big\|\,\Gamma_{R_X}^{\otimes n}\big) - D_{\max}^{\hat\epsilon'}\big(\rho_{X'}^{\otimes n}\,\big\|\,\Gamma_{X'}^{\otimes n}\big) + \log\frac{\hat\epsilon'''^2}{2+\hat\epsilon'''^2}\Big]\\ &\quad= D(\rho_{R_X}\,\|\,\Gamma_{R_X}) - D(\rho_{X'}\,\|\,\Gamma_{X'})\ .\end{aligned} \qquad ■$$

Equation (C.141) follows directly from (C.140), using the relations given by Proposition 25.

Appendix D: Robustness of battery states to smoothing

Because the battery system is a part of the physical implementation of the process, we may ask why it is not included in the definition of the smooth coherent relative entropy (C.1) in a way which would allow the physical implementation to fail to produce the appropriate output battery state with a small probability. Remarkably, there would have been no difference had we chosen to smooth the battery states as well. This follows from the following proposition, which asserts that optimization candidates which include smoothing on the battery states are in fact already included in the optimization in the definition above. This holds for the general battery states of the form $P_A\Gamma_A P_A/\operatorname{tr}(P_A\Gamma_A)$, for a projector P_A commuting with the Γ_A of the battery (see item (v) of Proposition 3).

Proposition 39 (Smoothing battery states)**.** *Let A, A' be quantum systems with corresponding Γ_A, $\Gamma_{A'}$. Let P_A, $P'_{A'}$ be projectors such that $[P_A,\Gamma_A] = 0$ and $[P_{A'},\Gamma_{A'}] = 0$, and let $\Phi_{XA\to X'A'}$ be a trace nonincreasing, completely positive map such that $\Phi_{XA\to X'A'}(\Gamma_X\otimes\Gamma_A) \leqslant \Gamma_{X'}\otimes\Gamma_{A'}$, and such that*

$$P\left[\Phi_{XA\to X'A}\Big(\sigma_{XR}\otimes\frac{P_A\Gamma_A P_A}{\operatorname{tr}P_A\Gamma_A}\Big)\,,\ \rho_{X'R}\otimes\frac{P'_{A'}\Gamma_{A'}P'_{A'}}{\operatorname{tr}P'_{A'}\Gamma_{A'}}\right] \leqslant \epsilon\ , \tag{D.1}$$

Then there exists a trace-nonincreasing, completely positive map $\mathcal{T}_{X\to X'}$ such both the following conditions hold:

$$P(\mathcal{T}_{X\to X'}(\sigma_{XR}),\rho_{X'R}) \leqslant \epsilon\ ; \tag{D.2a}$$

$$\mathcal{T}_{X\to X'}(\Gamma_X) \leqslant \frac{\operatorname{tr}(P'_{A'}\Gamma_{A'})}{\operatorname{tr}(P_A\Gamma_A)}\,\Gamma_{X'}\ . \tag{D.2b}$$

Proof of Proposition 39. Define, for any ω_X,

$$\mathcal{T}_{X\to X'}(\omega_X) = \operatorname{tr}_{A'}\left[P'_{A'}\,\Phi_{XA\to X'A'}\Big(\omega_X\otimes\frac{P_A\Gamma_A P_A}{\operatorname{tr}P_A\Gamma_A}\Big)\right]\ . \tag{D.3}$$

Then

$$\begin{aligned}\mathcal{T}_{X\to X'}(\sigma_{XR}) &= \operatorname{tr}_{A'}\left[P'_{A'}\,\Phi_{XA\to X'A'}\Big(\sigma_{XR}\otimes\frac{P_A\Gamma_A P_A}{\operatorname{tr}P_A\Gamma_A}\Big)\right]\\ &= \operatorname{tr}_{A'}\left[P'_{A'}\,\tilde\rho_{A'X'R}\right]\ ,\end{aligned} \tag{D.4}$$

where $\tilde\rho_{A'X'R} := \Phi_{XA\to X'A'}(\sigma_{XR}\otimes\frac{P_A\Gamma_A P_A}{\operatorname{tr}P_A\Gamma_A})$ satisfies $P(\tilde\rho_{A'X'R},\rho_{X'R}\otimes\frac{P'_{A'}\Gamma_{A'}P'_{A'}}{\operatorname{tr}P'_{A'}\Gamma_{A'}}) \leqslant \epsilon$ by assumption. Using the monotonicity of the purified distance [69] in particular under the trace-nonincreasing completely positive map $\operatorname{tr}\left[P'_{A'}(\cdot)\right]$, we have

$$P\big(\mathcal{T}_{X\to X'}(\sigma_{XR}),\rho_{X'R}\big) \leqslant \epsilon\ . \tag{D.5}$$

We also have

$$\begin{aligned}\mathcal{T}_{X\to X'}(\Gamma_X) &= \operatorname{tr}_{A'}\left[P'_{A'}\,\Phi_{XA\to X'A'}(\Gamma_X\otimes\frac{P_A\Gamma_A P_A}{\operatorname{tr}P_A\Gamma_A})\right]\\ &\leqslant \frac{1}{\operatorname{tr}P_A\Gamma_A}\cdot\operatorname{tr}_{A'}\left[P'_{A'}\,\Gamma_{X'}\otimes\Gamma_{A'}\right]\ ,\end{aligned} \tag{D.6}$$

using the fact that $P_A\Gamma_A P_A = \Gamma_A^{1/2}P_A\Gamma_A^{1/2} \leqslant \Gamma_A$ (because $[P_A,\Gamma_A] = 0$) and also with the fact that $\Phi_{XA\to X'A'}$ is Γ-sub-preserving. Then

$$\mathcal{T}_{X\to X'}(\Gamma_X) \leqslant \frac{\operatorname{tr}P'_{A'}\Gamma_{A'}}{\operatorname{tr}P_A\Gamma_A}\,\Gamma_{X'}\ , \tag{D.7}$$

which completes the proof. ■

This means that the processes which also allow "fuzziness" on the battery states are *de facto* already included in the optimization defining the smooth coherent relative entropy (C.1). This is formulated explicitly in the following corollary.

Corollary 40. *Let $\rho_{X'R_X}$ be a subnormalized state, let $\Gamma_X, \Gamma_{X'} \geqslant 0$ and let $\epsilon > 0$. Then*

$$\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) = \max_{A, A', P_A, P'_{A'}, \Phi_{XA\to X'A'}} -\log \frac{\operatorname{tr} P'_{A'}\Gamma_{A'}}{\operatorname{tr} P_A \Gamma_A} \,, \tag{D.8}$$

where the optimization is performed over all systems A, A', all operators Γ_A, $\Gamma_{A'}$, and all projectors P_A, $P'_{A'}$ such that $[P_A, \Gamma_A] = 0$ and $[P'_{A'}, \Gamma_{A'}] = 0$, for which there is a trace nonincreasing, completely positive map $\Phi_{XA\to X'A'}$ satisfying $\Phi_{XA\to X'A'}(\Gamma_X \otimes \Gamma_A) \leqslant \Gamma_{X'} \otimes \Gamma_{A'}$ as well as

$$P\left[\Phi_{XA\to X'A'}\left(\sigma_{XR} \otimes \frac{P_A\Gamma_A P_A}{\operatorname{tr} P_A\Gamma_A}\right), \rho_{X'R} \otimes \frac{P'_{A'}\Gamma_{A'}P'_{A'}}{\operatorname{tr} P'_{A'}\Gamma_{A'}}\right] \leqslant \epsilon \,. \tag{D.9}$$

Proof of Corollary 40. First, let A, A', P_A, $P'_{A'}$, Γ_A, $\Gamma_{A'}$ and $\Phi_{XA\to X'A'}$ satisfy the conditions of the maximization (D.8). Let $\mathcal{T}_{X\to X'}$ the mapping given by Proposition 39. Observe that $\big\|\Gamma_{X'}^{-1/2}\mathcal{T}_{X\to X'}(\Gamma_X)\,\Gamma_{X'}^{-1/2}\big\|_\infty \leqslant (\operatorname{tr} P'_{A'}\Gamma_{A'})/(\operatorname{tr} P_A\Gamma_A)$. Note also that $P(\mathcal{T}_{X\to X'}(\sigma_{XR}), \rho_{X'R_X}) \leqslant \epsilon$ as guaranteed by our previous use of Proposition 39. Hence, $\mathcal{T}_{X\to X'}$ is a valid candidate in the optimization given by Proposition 11 for $\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'})$. Hence

$$\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'}) \geqslant -\log \frac{\operatorname{tr} P'_{A'}\Gamma_{A'}}{\operatorname{tr} P_A\Gamma_A} \,. \tag{D.10}$$

To show that equality is achieved in (D.8), let $\mathcal{T}_{X\to X'}$ be a valid optimization candidate in (C.1) for $\hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_X, \Gamma_{X'})$ which achieves the optimal value $y = \hat{D}^{\epsilon}_{X\to X'}(\rho_{X'R_X} \,\|\, \Gamma_R, \Gamma_{X'}) = -\log\big\|\Gamma_{X'}^{-1/2}\mathcal{T}_{X\to X'}(\Gamma_X)\,\Gamma_{X'}^{-1/2}\big\|_\infty$, with $P(\mathcal{T}_{X\to X'}(\sigma_{XR_X}), \rho_{X'R_X}) \leqslant \epsilon$. Then $\mathcal{T}_{X\to X'}(\Gamma_X) \leqslant 2^{-y}\,\Gamma_{X'}$, and this mapping satisfies the conditions of item (i) of Proposition 3. Let $A = A'$ be a qubit system with $P_A = |0\rangle\langle 0|_A$, $P'_{A'} = |1\rangle\langle 1|_{A'}$, and $\Gamma_A = \Gamma_{A'} = g_0\,|0\rangle\langle 0|_A + g_1\,|1\rangle\langle 1|_A$, with $g_0/g_1 = 2^y$. In virtue of item (iii) of Proposition 3, there exists a trace-nonincreasing, completely positive map $\Phi_{XA\to X'A'}$ such that $\Phi_{XA\to X'A'}(\Gamma_X \otimes \Gamma_A) \leqslant \Gamma_{X'} \otimes \Gamma_{A'}$ and which satisfies $\Phi_{XA\to X'A'}((\cdot) \otimes |0\rangle\langle 0|_A) = \mathcal{T}_{X\to X'}(\cdot) \otimes |1\rangle\langle 1|_{A'}$. Then

$$\Phi_{XA\to X'A'}(\sigma_{XR_X} \otimes |0\rangle\langle 0|_A) = \mathcal{T}_{X\to X'}(\sigma_{XR_X}) \otimes |1\rangle\langle 1|_{A'} \,, \tag{D.11}$$

and hence

$$P(\Phi_{XA\to X'A'}(\sigma_{XR_X} \otimes |0\rangle\langle 0|_A), \rho_{X'R_X} \otimes |1\rangle\langle 1|_{A'}) = P(\mathcal{T}_{X\to X'}(\sigma_{XR_X}), \rho_{X'R_X}) \leqslant \epsilon \,. \tag{D.12}$$

Hence, all the conditions of the maximization (D.8) are satisfied, and the achieved value is indeed $-\log[(\operatorname{tr} P'_{A'}\Gamma_{A'})/(\operatorname{tr} P_A\Gamma_A)] = -\log(g_1/g_0) = y$. ∎

Appendix E: Technical Utilities

Lemma 41. *Let $A \geqslant 0$, $B \geqslant 0$ and let Π be the projector onto the support of A. Let $\mu > 0$. Define P as the projector onto the eigenspaces associated to nonnegative eigenvalues of the operator $(\mu A - B)$. Then there exists a constant c which is independent of μ such that*

$$\|\Pi - P\Pi P\|_\infty \leqslant \frac{c}{\mu} \,. \tag{E.1}$$

In particular,

$$\Pi \leqslant P + \frac{c}{\mu}\mathbb{1} \,. \tag{E.2}$$

Proof of Lemma 41. This lemma follows from a result of perturbation of matrix eigenspaces [110]. We'll consider the operators $A - \frac{1}{\mu}B$ and A. Let $Q = \mathbb{1} - P$ be the projector on the eigenspaces associated to the strictly negative eigenvalues of $A - \frac{1}{\mu}B$. Let $a_{\min} = \|A^{-1}\|_\infty^{-1}$ be the smallest nonzero eigenvalue of A. Recall that Π projects onto the eigenspaces of A associated to eigenvalues larger or equal to $a_{\min}$. We may now invoke [110, Theorem VII.3.1], which asserts that for any unitarily invariant norm $\|\cdot\|_\bullet$,

$$\|Q\Pi\|_\bullet \leqslant \frac{1}{\mu a_{\min}}\|QB\Pi\|_\bullet \leqslant \frac{1}{\mu a_{\min}}\|B\|_\bullet \,. \tag{E.3}$$

(The gap δ in [110, Theorem VII.3.1] is here the gap between 0 and $a_{\min}$.) In particular, we have $\|Q\Pi\|_\infty \leqslant (\mu a_{\min})^{-1}\,\|B\|_\infty$. We then have

$$\begin{aligned}\|\Pi - P\Pi P\|_\infty &\leqslant \|\Pi - P\Pi\|_\infty + \|P\Pi - P\Pi P\|_\infty \\ &\leqslant \|\Pi - P\Pi\|_\infty + \|P\|_\infty\|\Pi - \Pi P\|_\infty \\ &= 2\|\Pi - P\Pi\|_\infty = 2\|Q\Pi\|_\infty \leqslant \frac{c}{\mu} \,,\end{aligned} \tag{E.4}$$

with $c = 2\,(a_{\min})^{-1}\,\|B\|_\infty$. This implies (E.2) because

$$\Pi - P\Pi P \leqslant \frac{c}{\mu}\mathbb{1} \qquad \Rightarrow \qquad \Pi \leqslant P\Pi P + \frac{c}{\mu}\mathbb{1} \leqslant P + \frac{c}{\mu}\mathbb{1} \,.$$ ∎

Lemma 42. *Let ρ and σ be quantum states. The trace distance $D(\rho, \sigma)$ between ρ and σ can be written as the semidefinite program in terms of the variables $\Delta^\pm \geqslant 0$:*

$$\text{minimize}: \quad \frac{1}{2}\operatorname{tr}(\Delta^+ + \Delta^-) \tag{E.5a}$$

$$\text{subject to}: \quad \sigma = \rho + \Delta^+ - \Delta^- \,. \tag{E.5b}$$

Furthermore, $\operatorname{tr}\Delta^+ = \operatorname{tr}\Delta^- = D(\rho, \sigma)$ for the optimal solution. The dual to this program is an alternate expression of the same quantity, in terms of the Hermitian variable Z:

$$\text{maximize}: \quad \frac{1}{2}\operatorname{tr}(Z(\rho - \sigma)) \tag{E.6a}$$

$$\text{subject to}: \quad -\mathbb{1} \leqslant Z \leqslant \mathbb{1} \,. \tag{E.6b}$$

Proof of Lemma 42. Write $D(\rho, \sigma) = \frac{1}{2}\|\rho - \sigma\|_1$ and recall that for any Hermitian A, $\|A\|_1 = \operatorname{tr}|A|$. Choosing $\Delta_\pm \geqslant 0$ as the positive and negative parts of $\rho - \sigma$, i.e. such that $\rho - \sigma = \Delta + -\Delta_-$, yields feasible candidates for the primal problem and $\frac{1}{2}\operatorname{tr}(\Delta_+ + \Delta_-) = \frac{1}{2}\operatorname{tr}|\rho - \sigma| = D(\rho, \sigma)$. Now let $\Pi_\pm$ be the projectors onto the strictly positive and strictly negative parts of $\rho - \sigma$, respectively, and choose $Z = \Pi_+ - \Pi_-$. Observe that $\Pi_\pm(\rho - \sigma) = \pm\Delta_\pm$. Then $\frac{1}{2}\operatorname{tr}(Z(\rho - \sigma)) = \frac{1}{2}\operatorname{tr}(\Delta_+ + \Delta_-) = D(\rho, \sigma)$. We have exhibited primal and dual candidates achieving the value $D(\rho, \sigma)$, and hence this is the optimal solution of the semidefinite program. Furthermore (E.5b) implies that $\operatorname{tr}\Delta_+ = \operatorname{tr}\Delta_-$ and hence $\operatorname{tr}\Delta_+ = \operatorname{tr}\Delta_- = \frac{1}{2}\operatorname{tr}(\Delta_+ + \Delta_-) = D(\rho, \sigma)$. ∎

Lemma 43 (Gentle measurement lemma for the purified distance)**.** *Let $\rho \geqslant 0$ with $\operatorname{tr}\rho = 1$. Let $\epsilon \geqslant 0$. Let Π be a projector such that $\operatorname{tr}(\Pi\rho) \geqslant 1-\epsilon$. Then*

$$P\left(\rho, \frac{\Pi\rho\Pi}{\operatorname{tr}(\Pi\rho)}\right) \leqslant \frac{\sqrt{2\epsilon}}{\sqrt{1-\epsilon}} \,. \tag{E.7}$$

Proof of Lemma 43. Calculate

$$\begin{aligned} P^2\left(\rho, \frac{\Pi\rho\Pi}{\operatorname{tr}(\Pi\rho)}\right) &= 1 - F^2\left(\rho, \frac{\Pi\rho\Pi}{\operatorname{tr}(\Pi\rho)}\right) \\ &= \frac{1}{\operatorname{tr}(\Pi\rho)}\left[\operatorname{tr}(\Pi\rho) - F^2(\rho, \Pi\rho\Pi)\right] \\ &\leqslant \frac{1}{\operatorname{tr}(\Pi\rho)}\left[1 - F^2(\rho, \Pi\rho\Pi)\right] \\ &\leqslant \frac{1}{1-\epsilon} P^2(\rho, \Pi\rho\Pi)\,, \end{aligned} \tag{E.8}$$

noting that the generalized fidelity is $F(\rho,\sigma) = \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1$ as long as one of the states is normalized, and hence for $a > 0$ we have $F^2(\rho, a\sigma) = a\,F^2(\rho,\sigma)$ if $\operatorname{tr}\rho = 1$. Now, applying [111, Lemma 7], we have

$$P\left(\rho, \frac{\Pi\rho\Pi}{\operatorname{tr}(\Pi\rho)}\right) \leqslant \frac{\sqrt{2\epsilon-\epsilon^2}}{\sqrt{1-\epsilon}} = \frac{\sqrt{\epsilon\,(2-\epsilon)}}{\sqrt{1-\epsilon}} \leqslant \frac{\sqrt{2\epsilon}}{\sqrt{1-\epsilon}} \,.$$ ∎

Lemma 44 (Smoothing "part of" a state)**.** *Let ρ_{AB} be a bipartite normalized quantum state and let $\tilde\rho_A$ be a normalized quantum state such that $D(\tilde\rho_A, \rho_A) \leqslant \delta$. Then there exists a normalized quantum state $\hat\rho_{AB}$ such that $\operatorname{tr}_B \hat\rho_{AB} = \tilde\rho_A$, $\operatorname{tr}_A \hat\rho_{AB} = \rho_B$ and $P(\hat\rho_{AB}, \rho_{AB}) \leqslant 2\sqrt{2\delta}$.*

Proof of Lemma 44. Because $\tilde\rho_A$ and ρ_A are δ-close in trace distance, by Lemma 42 there exists $\Delta_A^\pm \geqslant 0$ such that $\operatorname{tr}\Delta_A^- = \operatorname{tr}\Delta_A^+ = D(\tilde\rho_A, \rho_A) \leqslant \delta$ and

$$\tilde\rho_A = \rho_A + \Delta_A^+ - \Delta_A^- \,. \tag{E.9}$$

Let $A = \tilde\rho_A + \Delta_A^- \geqslant 0$ and let $M_A = \tilde\rho_A^{1/2} A^{-1/2}$. Observe that $M_A^\dagger M_A = A^{-1/2}\,\tilde\rho_A\,A^{-1/2} \leqslant \mathbb{1}$ since $\tilde\rho_A \leqslant A$. Now define the completely positive map

$$\mathcal{M}_{A\to A}(\cdot) = M_A\,(\cdot)\,M_A^\dagger + \operatorname{tr}\left[(\mathbb{1} - M_A^\dagger M_A)(\cdot)\right]\xi_A \,, \tag{E.10}$$

with $\xi_A := (M_A\,\Delta_A^+\,M_A^\dagger)/\operatorname{tr}(M_A\,\Delta_A^+\,M_A^\dagger) \geqslant 0$ except if $\operatorname{tr}(M_A\,\Delta_A^+\,M_A^\dagger) = 0$, in which case we set $\xi_A := \mathbb{1}_A/|A|$. In any case $\operatorname{tr}\xi_A = 1$ and $\operatorname{tr}(M_A\,\Delta_A^+\,M_A^\dagger)\,\xi_A = M_A\,\Delta_A^+\,M_A^\dagger$. The mapping $\mathcal{M}_{A\to A}$ is trace preserving:

$$\mathcal{M}^\dagger(\mathbb{1}_A) = M_A^\dagger\,M_A + (\mathbb{1} - M_A^\dagger M_A)\,\operatorname{tr}\xi_A = \mathbb{1}_A \,. \tag{E.11}$$

We now show that $\mathcal{M}_{A\to A}(\rho_A) = \tilde\rho_A$. On one hand, using $A = \tilde\rho_A + \Delta_A^- = \rho_A + \Delta_A^+$, we have

$$M_A\,\rho_A\,M_A^\dagger = M_A\,A\,M_A^\dagger - M_A\,\Delta_A^+\,M_A^\dagger = \tilde\rho_A - M_A\,\Delta_A^+\,M_A^\dagger \,. \tag{E.12}$$

while noting that ρ_A lies within the support of A since $A = \rho_A + \Delta_A^+$. We deduce that $\operatorname{tr}(M_A\,\rho_A\,M_A^\dagger) = 1 - \operatorname{tr}(M_A\,\Delta_A^+\,M_A^\dagger)$. On the other hand,

$$\begin{aligned} \operatorname{tr}\left[(\mathbb{1} - M_A^\dagger M_A)\,\rho_A\right]\xi_A &= \left(1 - \operatorname{tr}(M_A\,\rho_A\,M_A^\dagger)\right)\xi_A \\ &= \operatorname{tr}(M_A\,\Delta_A^+\,M_A^\dagger)\,\xi_A \\ &= M_A\,\Delta_A^+\,M_A^\dagger \,, \end{aligned} \tag{E.13}$$

and hence, combining (E.12) with (E.13)

$$\mathcal{M}_{A\to A}(\rho_A) = M_A\,\rho_A\,M_A^\dagger + \operatorname{tr}\left[(\mathbb{1} - M_A^\dagger M_A)\,\rho_A\right]\xi_A = \tilde\rho_A \,. \tag{E.14}$$

Define now the state $\hat\rho_{AB}$ as

$$\hat\rho_{AB} = \mathcal{M}_{A\to A}\left[\rho_{AB}\right] \tag{E.15}$$

where the identity mapping is understood on system B. By properties of quantum channels the state on B is preserved, i.e. $\operatorname{tr}_A\hat\rho_{AB} = \rho_B$ (and in particular we have $\operatorname{tr}\hat\rho_{AB} = 1$), and we showed above that $\operatorname{tr}_B\hat\rho_{AB} = \tilde\rho_A$. It remains to see that $\hat\rho_{AB}$ and ρ_{AB} are close in purified distance. Let $|\rho\rangle_{ABC}$ be a purification of ρ_{AB}. Apply [106, Lemma A.4]—itself a reformulation of [112, Lemma 15]—with $\rho_{\text{Lem A.4}} = \rho_A$, $\sigma_{\text{Lem A.4}} = \tilde\rho_A$, $\Delta_{\text{Lem A.4}} = \Delta_A^-$, $G_{\text{Lem A.4}} = M_A$ and $|\psi_{\text{Lem A.4}}\rangle = |\rho\rangle_{ABC}$ to obtain

$$P(M_A\rho_{ABC}M_A^\dagger, \rho_{ABC}) \leqslant \sqrt{\left(2 - \operatorname{tr}\Delta_A^-\right)\operatorname{tr}\Delta_A^-} \leqslant \sqrt{2\delta} \,. \tag{E.16}$$

This distance can only decrease if we trace out the system C, and thus $P(M_A\rho_{AB}M_A^\dagger, \rho_{AB}) \leqslant \sqrt{2\delta}$. On the other hand, we have by definition

$$\hat\rho_{AB} = M_A\,\rho_{AB}\,M_A^\dagger + \Delta'_{AB} \,, \tag{E.17}$$

with $\Delta'_{AB} = \operatorname{tr}_A[(\mathbb{1}_A - M_A^\dagger M_A)\,\rho_{AB}] \otimes \xi_A \geqslant 0$. Calculate $\operatorname{tr}\Delta'_{AB} = \operatorname{tr}\left[(\mathbb{1}_A - M_A^\dagger M_A)\,\rho_A\right] = \operatorname{tr}(M_A\,\Delta_A^+\,M_A^\dagger) \leqslant \operatorname{tr}\Delta_A^+ \leqslant \delta$, and hence $D(M_A\,\rho_{AB}\,M_A^\dagger, \hat\rho_{AB}) \leqslant \delta$. Finally, by triangle inequality and using $P(\rho,\rho') \leqslant \sqrt{2D(\rho,\rho')}$,

$$P(\hat\rho_{AB}, \rho_{AB}) \leqslant P(\hat\rho_{AB}, M_A\,\rho_{AB}\,M_A^\dagger) + P(M_A\,\rho_{AB}\,M_A^\dagger, \rho_{AB}) \leqslant 2\sqrt{2\delta} \,.$$ ∎

Lemma 45. *Let σ_X, $\hat\sigma_X$ be two states. Consider another system $R \simeq X$. Then*

$$P(\sigma_X^{1/2}\,\Phi_{X:R}\,\sigma_X^{1/2}, \hat\sigma_X^{1/2}\,\Phi_{X:R}\,\hat\sigma_X^{1/2}) \leqslant 2\,\sqrt{D(\sigma_X, \hat\sigma_X)} \,. \tag{E.18}$$

Proof of Lemma 45. Let $\epsilon = D(\sigma_X, \hat\sigma_X)$. Using the properties of the trace distance, let $\Delta_X^\pm \geqslant 0$ satisfy $\hat\sigma_X = \sigma_X + \Delta_X^+ - \Delta_X^-$ with $\operatorname{tr}\Delta^+ = \operatorname{tr}\Delta^- = \epsilon$. Let $|\psi\rangle = (1+\epsilon)^{-1/2}\,(\sigma_X + \Delta_X^+)^{1/2}|\Phi\rangle_{X:R} = (1+\epsilon)^{-1/2}\,(\hat\sigma_X + \Delta_X^-)^{1/2}|\Phi\rangle_{X:R}$, noting that $\langle\psi\,|\,\psi\rangle = \operatorname{tr}(\sigma_X + \Delta_X^+)/(1-\epsilon) = 1$. For any two pure states $|\phi\rangle$, $|\chi\rangle$ we know that $P(|\phi\rangle\langle\phi|, |\chi\rangle\langle\chi|) = \left(1 - |\langle\phi\,|\,\chi\rangle|^2\right)^{1/2}$. Our strategy for proving the claim is the following: We show that both

$$|\langle\psi_{XR}\,|\,\sigma_X^{1/2}\,|\Phi_{X:R}\rangle| \geqslant (1+\epsilon)^{-1/2} \,; \tag{E.19a}$$

$$|\langle\psi_{XR}\,|\,\hat\sigma_X^{1/2}\,|\Phi_{X:R}\rangle| \geqslant (1+\epsilon)^{-1/2} \,, \tag{E.19b}$$

and the claim will then follow by triangle inequality for the purified distance: $P(\sigma_X^{1/2}\,\Phi_{X:R}\,\sigma_X^{1/2}, \hat\sigma_X^{1/2}\,\Phi_{X:R}\,\hat\sigma_X^{1/2}) \leqslant P(\sigma_X^{1/2}\,\Phi_{X:R}\,\sigma_X^{1/2}, |\psi\rangle\langle\psi|) + P(\hat\sigma_X^{1/2}\,\Phi_{X:R}\,\hat\sigma_X^{1/2}, |\psi\rangle\langle\psi|) \leqslant 2\sqrt{1 - 1/(1+\epsilon)} \leqslant 2\sqrt{\epsilon/(1+\epsilon)} \leqslant 2\sqrt{\epsilon}$. It remains to show the properties (E.19). We have $\langle\psi_{XR}\,|\,\sigma_X^{1/2}\,|\Phi_{X:R}\rangle = \langle\Phi_{X:R}\,|(\sigma_X^{1/2} + \Delta_X^+)^{1/2}\sigma_X^{1/2}\,|\Phi_{X:R}\rangle/\sqrt{1+\epsilon} = \operatorname{tr}\left[(\sigma_X + \Delta_X^+)^{1/2}\sigma_X^{1/2}\right]/\sqrt{1+\epsilon} \geqslant 1/\sqrt{1+\epsilon}$, noting that $(\sigma_X + \Delta_X^+)^{1/2} \geqslant (\sigma_X)^{1/2}$, and hence $|\langle\psi_{XR}\,|\,\sigma_X^{1/2}\,|\Phi_{X:R}\rangle| \geqslant (1+\epsilon)^{-1/2}$. Similarly, $\langle\psi_{XR}\,|\,\hat\sigma_X^{1/2}\,|\Phi_{X:R}\rangle = \langle\Phi_{X:R}\,|(\hat\sigma_X^{1/2} + \Delta_X^-)^{1/2}\hat\sigma_X^{1/2}\,|\Phi_{X:R}\rangle/\sqrt{1+\epsilon} \geqslant (1+\epsilon)^{-1/2}$. ∎

Lemma 46 (Continuity of the relative entropy in its first argument)**.** *Let $\Gamma \geqslant 0$. Let ρ, σ lie within the support of Γ. Assume that $D(\rho,\sigma) \leqslant \epsilon$. Then*

$$\begin{aligned} &|D(\rho\,\|\,\Gamma) - D(\sigma\,\|\,\Gamma)| \\ &\qquad \leqslant \epsilon\log(\operatorname{rank}\Gamma - 1) + h(\epsilon) + \epsilon\,\|\log\Gamma\|_\infty \,, \end{aligned} \tag{E.20}$$

where $h(\epsilon) = -\epsilon\log\epsilon - (1-\epsilon)\log(1-\epsilon)$ is the binary entropy.

Proof of Lemma 46. First, write

$$D(\rho\,\|\,\Gamma) = \operatorname{tr}\left[\rho\log\rho - \rho\log\Gamma\right] = -H(\rho) - \operatorname{tr}\left[\rho\log\Gamma\right] , \tag{E.21}$$

and so

$$|D(\rho \,\|\, \Gamma) - D(\sigma \,\|\, \Gamma)| \leqslant |H(\sigma) - H(\rho)| + |\mathrm{tr}\,[\sigma \log \Gamma] - \mathrm{tr}\,[\rho \log \Gamma]| \ . \tag{E.22}$$

Using the continuity bound of Audenaert [113], we have

$$|H(\rho) - H(\sigma)| \leqslant \epsilon \ \log(\mathrm{rank}\,\Gamma - 1) + h(\epsilon)\,, \tag{E.23}$$

where the states ρ and σ can be seen as living in a subspace of the full Hilbert space of dimension at most Γ (because they must both lie within the support of Γ), and where $h(\epsilon) = -\epsilon \ln \epsilon - (1-\epsilon)\ln(1-\epsilon)$ is the binary entropy. On the other hand,

$$\begin{aligned} \mathrm{tr}\,\rho \log \Gamma - \mathrm{tr}\,\sigma \log \Gamma &= \|\log \Gamma\|_\infty \ \mathrm{tr}\left[(\rho - \sigma)\frac{\log \Gamma}{\|\log \Gamma\|_\infty}\right] \\ &\leqslant \|\log \Gamma\|_\infty \, D(\rho, \sigma)\,, \end{aligned}$$

as $\log \Gamma / \|\log \Gamma\|_\infty$ is a valid candidate for Z in Lemma 42. Inverting the roles of ρ and σ in the equation above we finally obtain:

$$|\mathrm{tr}\,\rho \log \Gamma - \mathrm{tr}\,\sigma \log \Gamma| \leqslant \|\log \Gamma\|_\infty \, D(\rho, \sigma) \leqslant \|\log \Gamma\|_\infty \cdot \epsilon \ . \qquad \blacksquare$$

[1] R. Landauer, "Irreversibility and heat generation in the computing process," IBM Journal of Research and Development **5**, 183–191 (1961).

[2] C. H. Bennett, "The thermodynamics of computation—a review," International Journal of Theoretical Physics **21**, 905–940 (1982).

[3] C. H. Bennett, "Notes on Landauer's principle, reversible computation and Maxwell's demon," Studies in History and Philosophy of Modern Physics **34**, 501–510 (2003), arXiv:physics/0210005.

[4] L. Szilard, "Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen," Zeitschrift für Physik **53**, 840–856 (1929).

[5] H. Scovil and E. Schulz-DuBois, "Three-level masers as heat engines," Physical Review Letters **2**, 262–263 (1959).

[6] J. Geusic, E. Schulz-DuBois, and H. Scovil, "Quantum equivalent of the carnot cycle," Physical Review **156**, 343–351 (1967).

[7] R. Alicki, "On the entropy production for the davies model of heat conduction," Journal of Statistical Physics **20**, 671–677 (1979).

[8] E. Geva and R. Kosloff, "On the classical limit of quantum thermodynamics in finite time," The Journal of Chemical Physics **97**, 4398 (1992).

[9] N. Linden, S. Popescu, and P. Skrzypczyk, "How small can thermal machines be? the smallest possible refrigerator," Physical Review Letters **105**, 130401 (2010), arXiv:0908.2076.

[10] O. Abah, J. Roßnagel, G. Jacob, S. Deffner, F. Schmidt-Kaler, K. Singer, and E. Lutz, "Single-ion heat engine at maximum power," Physical Review Letters **109**, 203006 (2012), arXiv:1205.1362.

[11] A. Roulet, S. Nimmrichter, J. M. Arrazola, S. Seah, and V. Scarani, "Autonomous rotor heat engine," Physical Review E **95**, 062131 (2017), arXiv:1609.06011.

[12] E. T. Jaynes, "Information theory and statistical mechanics," Physical Review **106**, 620–630 (1957).

[13] E. T. Jaynes, "Information theory and statistical mechanics. II," Physical Review **108**, 171–190 (1957).

[14] K. Shizume, "Heat generation required by information erasure," Physical Review E **52**, 3495–3499 (1995).

[15] B. Piechocinska, "Information erasure," Physical Review A **61**, 062314 (2000).

[16] P. Hänggi and F. Marchesoni, "Artificial brownian motors: Controlling transport on the nanoscale," Reviews of Modern Physics **81**, 387–442 (2009).

[17] J. Anders, S. Shabbir, S. Hilt, and E. Lutz, "Landauer's principle in the quantum domain," Electronic Proceedings in Theoretical Computer Science **26**, 13–18 (2010), arXiv:1006.1420.

[18] T. Sagawa and M. Ueda, "Fluctuation theorem with information exchange: Role of correlations in stochastic thermodynamics," Physical Review Letters **109**, 180602 (2012), arXiv:1206.2479.

[19] J. Goold, M. Paternostro, and K. Modi, "A non-equilibrium quantum Landauer principle," Physical Review Letters **114**, 060602 (2015), arXiv:1402.4499.

[20] S. Toyabe, T. Sagawa, M. Ueda, E. Muneyuki, and M. Sano, "Experimental demonstration of information-to-energy conversion and validation of the generalized jarzynski equality," Nature Physics **6**, 988–992 (2010), arXiv:1009.5287.

[21] J. Baugh, O. Moussa, C. A. Ryan, A. Nayak, and R. Laflamme, "Experimental implementation of heat-bath algorithmic cooling using solid-state nuclear magnetic resonance," Nature **438**, 470–3 (2005), arXiv:quant-ph/0512024.

[22] D. Janzing, P. Wocjan, R. Zeier, R. Geiss, and T. Beth, "Thermodynamic cost of reliability and low temperatures: Tightening Landauer's principle and the second law," International Journal of Theoretical Physics **39**, 2717–2753 (2000), arXiv:quant-ph/0002048.

[23] M. Horodecki, P. Horodecki, and J. Oppenheim, "Reversible transformations from pure to mixed states and the unique measure of information," Physical Review A **67**, 062104 (2003), arXiv:quant-ph/0212019.

[24] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, "Resource theory of quantum states out of thermal equilibrium," Physical Review Letters **111**, 250404 (2013), arXiv:1111.3882.

[25] M. Horodecki and J. Oppenheim, "Fundamental limitations for quantum and nanoscale thermodynamics," Nature Communications **4**, 2059 (2013), arXiv:1111.3834.

[26] P. Skrzypczyk, A. J. Short, and S. Popescu, "Work extraction and thermodynamics for individual quantum systems." Nature communications **5**, 4185 (2014), arXiv:1307.1558.

[27] G. Gour, M. P. Müller, V. Narasimhachar, R. W. Spekkens, and N. Y. Halpern, "The resource theory of informational nonequilibrium in thermodynamics," Physics Reports (2015), 10.1016/j.physrep.2015.04.003, arXiv:1309.6586.

[28] F. Brandão, M. Horodecki, N. Ng, J. Oppenheim, and S. Wehner, "The second laws of quantum thermodynamics," Proceedings of the National Academy of Sciences **112**, 201411728 (2015), arXiv:1305.5278.

[29] F. G. S. L. Brandão and G. Gour, "Reversible framework for quantum resource theories," Physical Review Letters **115**, 070503 (2015), arXiv:1502.03149.

[30] O. C. O. Dahlsten, R. Renner, E. Rieper, and V. Vedral, "Inadequacy of von Neumann entropy for characterizing extractable work," New Journal of Physics **13**, 53015 (2011), arXiv:0908.0424.

[31] J. Åberg, "Truly work-like work extraction via a single-shot analysis," Nature Communications **4**, 1925 (2013), arXiv:1110.6121.

[32] D. Egloff, O. C. O. Dahlsten, R. Renner, and V. Vedral, "A measure of majorization emerging from single-shot statistical mechanics," New Journal of Physics **17**, 073001 (2015), arXiv:1207.0434.

[33] L. del Rio, J. Åberg, R. Renner, O. Dahlsten, and V. Vedral, "The thermodynamic meaning of negative entropy," Nature **474**, 61–63 (2011), arXiv:1009.1630.

[34] L. del Rio, A. Hutter, R. Renner, and S. Wehner, "Relative thermal-

ization," Physical Review E **94**, 022104 (2016), arXiv:1401.7997.

[35] S. Barnett and J. Vaccaro, "Beyond Landauer erasure," Entropy **15**, 4956–4968 (2013), arXiv:1310.7821.

[36] N. Yunger Halpern and J. M. Renes, "Beyond heat baths: Generalized resource theories for small-scale thermodynamics," Physical Review E **93**, 022126 (2016), arXiv:1409.3998.

[37] M. Perarnau-Llobet, A. Riera, R. Gallego, H. Wilming, and J. Eisert, "Work and entropy production in generalised Gibbs ensembles," New Journal of Physics **18**, 123035 (2016), arXiv:1512.03823.

[38] Y. Guryanova, S. Popescu, A. J. Short, R. Silva, and P. Skrzypczyk, "Thermodynamics of quantum systems with multiple conserved quantities," Nature Communications **7**, 12049 (2016), arXiv:1512.01190.

[39] N. Yunger Halpern, P. Faist, J. Oppenheim, and A. Winter, "Microcanonical and resource-theoretic derivations of the thermal state of a quantum system with noncommuting charges," Nature Communications **7**, 12051 (2016), arXiv:1512.01189.

[40] D. Reeb and M. M. Wolf, "An improved Landauer principle with finite-size corrections," New Journal of Physics **16**, 103011 (2014), arXiv:1306.4352.

[41] J. G. Richens, Á. M. Alhambra, and L. Masanes, "Finite-bath corrections to the second law of thermodynamics," (2017), arXiv:1702.03357.

[42] J. Åberg, "Catalytic coherence," Physical Review Letters **113**, 150402 (2014), arXiv:1304.1060.

[43] P. Ćwikliński, M. Studziński, M. Horodecki, and J. Oppenheim, "Limitations on the evolution of quantum coherences: Towards fully quantum second laws of thermodynamics," Physical Review Letters **115**, 210403 (2015), arXiv:1405.5029.

[44] N. H. Y. Ng, L. Mančinska, C. Cirstoiu, J. Eisert, and S. Wehner, "Limits to catalysis in quantum thermodynamics," New Journal of Physics **17**, 085004 (2015), arXiv:1405.3039.

[45] M. Lostaglio, K. Korzekwa, D. Jennings, and T. Rudolph, "Quantum coherence, time-translation symmetry, and thermodynamics," Physical Review X **5**, 021001 (2015), arXiv:1410.4572.

[46] M. Lostaglio, D. Jennings, and T. Rudolph, "Description of quantum coherence in thermodynamic processes requires constraints beyond free energy," Nature communications **6**, 6383 (2015), arXiv:1405.2188.

[47] K. Korzekwa, M. Lostaglio, J. Oppenheim, and D. Jennings, "The extraction of work from quantum coherence," New Journal of Physics **18**, 023045 (2016), arXiv:1506.07875.

[48] J. Oppenheim, M. Horodecki, P. Horodecki, and R. Horodecki, "Thermodynamical approach to quantifying quantum correlations," Physical Review Letters **89**, 180402 (2002), arXiv:quant-ph/0112074.

[49] D. E. Bruschi, M. Perarnau-Llobet, N. Friis, K. V. Hovhannisyan, and M. Huber, "Thermodynamics of creating correlations: Limitations and optimal protocols," Physical Review E **91**, 032118 (2015), arXiv:1409.4647.

[50] C. Majenz, M. Berta, F. Dupuis, R. Renner, and M. Christandl, "Catalytic decoupling of quantum information," Physical Review Letters **118**, 080503 (2017), arXiv:1605.00514.

[51] M. Lostaglio, M. P. Müller, and M. Pastena, "Stochastic independence as a resource in small-scale thermodynamics," Physical Review Letters **115**, 150402 (2015), arXiv:1409.3258.

[52] R. Gallego, A. Riera, and J. Eisert, "Thermal machines beyond the weak coupling regime," New Journal of Physics **16**, 125009 (2014), arXiv:1310.8349.

[53] M. P. Woods, N. Ng, and S. Wehner, "The maximum efficiency of nano heat engines depends on more than temperature," (2015), arXiv:1506.02322.

[54] M. Hayashi and H. Tajima, "Measurement-based formulation of quantum heat engines," Physical Review A **95**, 032132 (2017), arXiv:1504.06150.

[55] K. Ito and M. Hayashi, "Optimal performance of generalized heat engines with finite-size baths of arbitrary multiple conserved quantities beyond I.I.d. scaling," (2016), arXiv:1612.04047.

[56] J. Åberg, "Fully quantum fluctuation theorems," (2016), arXiv:1601.01302.

[57] Á. M. Alhambra, L. Masanes, J. Oppenheim, and C. Perry, "Fluctuating work: From quantum thermodynamical identities to a second law equality," Physical Review X **6**, 041017 (2016), arXiv:1601.05799.

[58] Á. M. Alhambra, S. Wehner, M. M. Wilde, and M. P. Woods, "Work and reversibility in quantum thermodynamics," (2015), arXiv:1506.08145.

[59] J. Goold, M. Huber, A. Riera, L. del Rio, and P. Skrzypczyk, "The role of quantum information in thermodynamics—a topical review," Journal of Physics A: Mathematical and Theoretical **49**, 143001 (2016), arXiv:1505.07835.

[60] P. Faist, F. Dupuis, J. Oppenheim, and R. Renner, "The minimal work cost of information processing," Nature Communications **6**, 7669 (2015), arXiv:1211.1037.

[61] R. Renner, *Security of Quantum Key Distribution*, Ph.D. thesis, ETH Zürich (2005), arXiv:quant-ph/0512258.

[62] N. Datta, "Min- and max-relative entropies and a new entanglement monotone," IEEE Transactions on Information Theory **55**, 2816–2826 (2009), arXiv:0803.2770.

[63] M. Tomamichel, *Quantum Information Processing with Finite Resources*, SpringerBriefs in Mathematical Physics, Vol. 5 (Springer International Publishing, Cham, 2016) arXiv:1504.00233.

[64] P. Faist, *Quantum Coarse-Graining: An Information-Theoretic Approach to Thermodynamics*, Ph.D. thesis, ETH Zurich (2016), arXiv:1607.03104.

[65] R. P. Feynman, *Lectures on Computation*, edited by Anthony J. G. Hey and Robin W. Allen (Westview Press, 1996).

[66] B. Coecke, T. Fritz, and R. W. Spekkens, "A mathematical theory of resources," Information and Computation **250**, 59–86 (2016), arXiv:1409.5531.

[67] L. Kraemer and L. del Rio, "Currencies in resource theories," (2016), arXiv:1605.01064.

[68] Á. M. Alhambra, J. Oppenheim, and C. Perry, "Fluctuating states: What is the probability of a thermodynamical transition?" Physical Review X **6**, 041016 (2016), arXiv:1504.00020.

[69] M. Tomamichel, *A Framework for Non-Asymptotic Quantum Information Theory*, Ph.D. thesis, ETH Zurich (2012), arXiv:1203.2142.

[70] M. Tomamichel, R. Colbeck, and R. Renner, "Duality between smooth min- and max-entropies," IEEE Transactions on Information Theory **56**, 4674–4681 (2010), arXiv:0907.5238.

[71] R. König, R. Renner, and C. Schaffner, "The operational meaning of min- and max-entropy," IEEE Transactions on Information Theory **55**, 4337–4347 (2009), arXiv:0807.1338.

[72] M. Tomamichel, C. Schaffner, A. Smith, and R. Renner, "Leftover hashing against quantum side information," IEEE Transactions on Information Theory **57**, 5524–5535 (2011), arXiv:1002.2436.

[73] F. Buscemi and N. Datta, "The quantum capacity of channels with arbitrarily correlated noise," IEEE Transactions on Information Theory **56**, 1447–1460 (2010), arXiv:0902.0158.

[74] F. G. S. L. Brandão and N. Datta, "One-shot rates for entanglement manipulation under non-entangling maps," IEEE Transactions on Information Theory **57**, 1754–1760 (2011), arXiv:0905.2673.

[75] M. Tomamichel and M. Hayashi, "A hierarchy of information quantities for finite block length analysis of quantum tasks," IEEE Transactions on Information Theory **59**, 7693–7710 (2013), arXiv:1208.1478.

[76] L. Wang and R. Renner, "One-shot classical-quantum capacity and hypothesis testing," Physical Review Letters **108**, 200501 (2012), arXiv:1007.5456.

[77] W. Matthews and S. Wehner, "Finite blocklength converse bounds for quantum channels," IEEE Transactions on Information Theory

60, 7317–7329 (2014), arXiv:1210.4722.

[78] M. Mosonyi and T. Ogawa, "Quantum hypothesis testing and the operational interpretation of the quantum Rényi relative entropies," Communications in Mathematical Physics **334**, 1617–1648 (2014), arXiv:1309.3228.

[79] R. Gallego, J. Eisert, and H. Wilming, "Thermodynamic work from operational principles," New Journal of Physics **18**, 103017 (2016), arXiv:1504.05056.

[80] E. T. Jaynes, "The Gibbs paradox," in *Maximum Entropy and Bayesian Methods*, Fundamental Theories of Physics, Vol. 50, edited by C. R. Smith, G. Erickson, and P. O. Neudorfer (Springer Netherlands, 1992) pp. 1–21.

[81] D. Petz, "Sufficient subalgebras and the relative entropy of states of a von Neumann algebra," Communications in Mathematical Physics **105**, 123–131 (1986).

[82] D. Petz, "Sufficiency of channels over von Neumann algebras," The Quarterly Journal of Mathematics **39**, 97–108 (1988).

[83] D. Petz, "Monotonicity of quantum relative entropy revisited," Reviews in Mathematical Physics **15**, 79–91 (2003), arXiv:quant-ph/0209053.

[84] M. M. Wilde, "Recoverability in quantum information theory," Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science **471**, 20150338 (2015), arXiv:1505.04661.

[85] S. Beigi, N. Datta, and F. Leditzky, "Decoding quantum information via the Petz recovery map," Journal of Mathematical Physics **57**, 082203 (2016), arXiv:1504.04449.

[86] R. Giles, *Mathematical Foundations of Thermodynamics* (Pergamon Press, 1964).

[87] E. P. Gyftopoulos and G. P. Beretta, *Thermodynamics: Foundations and Applications* (Dover Publications, 2005).

[88] E. H. Lieb and J. Yngvason, "The physics and mathematics of the second law of thermodynamics," Physics Reports **310**, 1–96 (1999), arXiv:cond-mat/9708200.

[89] E. H. Lieb and J. Yngvason, "A guide to entropy and the second law of thermodynamics," in *Statistical Mechanics*, edited by Bruno Nachtergaele, Jan Philip Solovej, and Jakob Yngvason (Springer Berlin Heidelberg, 2004) pp. 353–363, arXiv:math-ph/9805005.

[90] M. Weilenmann, L. Kraemer, P. Faist, and R. Renner, "Axiomatic relation between thermodynamic and information-theoretic entropies," Physical Review Letters **117**, 260601 (2016), arXiv:1501.06920.

[91] G. P. Beretta and E. Zanchini, "Rigorous and general definition of thermodynamic entropy," in *Thermodynamics*, edited by Mizutani Tadashi (InTech, 2011) Chap. 2, pp. 23–50, arXiv:1010.0813.

[92] O. Fawzi and R. Renner, "Quantum conditional mutual information and approximate markov chains," Communications in Mathematical Physics **340**, 575–611 (2015), arXiv:1410.0664.

[93] M. Berta, K. P. Seshadreesan, and M. M. Wilde, "Rényi generalizations of the conditional quantum mutual information," Journal of Mathematical Physics **56**, 022205 (2015), arXiv:1403.6102.

[94] N. Datta and M. M. Wilde, "Quantum markov chains, sufficiency of quantum channels, and Rényi information measures," Journal of Physics A: Mathematical and Theoretical **48**, 505301 (2015), arXiv:1501.05636.

[95] F. Buscemi, S. Das, and M. M. Wilde, "Approximate reversibility in the context of entropy gain, information gain, and complete positivity," Physical Review A **93**, 062314 (2016), arXiv:1601.01207.

[96] D. Sutter, M. Tomamichel, and A. W. Harrow, "Strengthened monotonicity of relative entropy via pinched Petz recovery map," IEEE Transactions on Information Theory **62**, 2907–2913 (2016), arXiv:1507.00303.

[97] D. Sutter, O. Fawzi, and R. Renner, "Universal recovery map for approximate markov chains," Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science **472**, 20150623 (2016), arXiv:1504.07251.

[98] P. Faist, J. Oppenheim, and R. Renner, "Gibbs-preserving maps outperform thermal operations in the quantum regime," New Journal of Physics **17**, 043003 (2015), arXiv:1406.3618.

[99] G. Vidal and J. Cirac, "Catalysis in nonlocal quantum operations," Physical Review Letters **88**, 167903 (2002), arXiv:quant-ph/0108077.

[100] R. van der Meer, N. H. Y. Ng, and S. Wehner, "Smoothed generalized free energies for thermodynamics," (2017), arXiv:1706.03193.

[101] M. Horodecki, J. Oppenheim, and C. Sparaciari, "Approximate majorization," (2017), arXiv:1706.05264.

[102] I. Bengtsson and K. Zyczkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, Cambridge, UK, 2006).

[103] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).

[104] J. Watrous, "Semidefinite programs for completely bounded norms," Theory of Computing **5**, 217–238 (2009), arXiv:0901.4709.

[105] J. Watrous, "Lecture notes on the theory of quantum information," (2011).

[106] F. Dupuis, L. Kraemer, Ph. Faist, J. M. Renes, and R. Renner, "Generalized entropies," in *XVIIth International Congress on Mathematical Physics* (2013) pp. 134–153, arXiv:1211.3141.

[107] A. Vitanov, F. Dupuis, M. Tomamichel, and R. Renner, "Chain rules for smooth min- and max-entropies," IEEE Transactions on Information Theory **59**, 2603–2612 (2013), arXiv:1205.5231.

[108] F. Hiai and D. Petz, "The proper formula for relative entropy and its asymptotics in quantum probability," Communications in Mathematical Physics **143**, 99–114 (1991).

[109] H. Nagaoka and T. Ogawa, "Strong converse and stein's lemma in quantum hypothesis testing," IEEE Transactions on Information Theory **46**, 2428–2433 (2000), arXiv:quant-ph/9906090.

[110] R. Bhatia, *Matrix Analysis*, Graduate Texts in Mathematics (Springer, New York, 1997).

[111] M. Berta, M. Christandl, R. Colbeck, J. M. Renes, and R. Renner, "The uncertainty principle in the presence of quantum memory," Nature Physics **6**, 659–662 (2010), arXiv:0909.0950.

[112] M. Tomamichel, R. Colbeck, and R. Renner, "A fully quantum asymptotic equipartition property," IEEE Transactions on Information Theory **55**, 5840–5847 (2009), arXiv:0811.1221.

[113] K. M. R. Audenaert, "A sharp continuity estimate for the von Neumann entropy," Journal of Physics A: Mathematical and Theoretical **40**, 8127–8136 (2007), arXiv:quant-ph/0610146.